



Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

**ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA AGRONÓMICA Y
BIOCIENCIAS**

***NEKAZARITZAKO INGENIARITZAKO ETA BIOZIENTZIETAKO GOI
MAILAKO ESKOLA TEKNIKOA***

*PUESTA A PUNTO DE UNA TÉCNICA DE IDENTIFICACIÓN
DE VARIANTES DE KAPPA-CASEÍNA EN LA RAZA BOVINA FRISONA*

presentado por

Leire Azparren Dominguez

GRADO EN INNOVACIÓN DE PROCESOS Y PRODUCTOS ALIMENTARIOS

GRADUA ELIKAGAI PROZESU ETA PRODUKTUEN BERRIKUNTZAN

Septiembre, 2021

ÍNDICE

| | |
|--|----|
| ABSTRACT | 5 |
| RESUMEN | 5 |
| 1. INTRODUCCIÓN | 6 |
| 1.1. La leche y los derivados lácteos | 6 |
| 1.2. Las proteínas lácteas | 7 |
| 1.3. La Kappa-caseína | 9 |
| 1.4. Evolución del proceso de selección genética del ganado bovino y las propiedades reológicas de la leche | 11 |
| 1.5. Técnicas de genotipado del ganado bovino empleadas para la mejora de las propiedades reológicas de la leche | 14 |
| 1.6. Técnica PCR cuantitativa en tiempo real o qPCR | 16 |
| 2. OBJETIVOS | 20 |
| 3. MATERIALES Y MÉTODOS | 21 |
| 3.1. Obtención de las muestras | 21 |
| 3.2. Extracción de ADN a partir de células somáticas de leche y pelo | 21 |
| 3.2.1. Metodología empleada para la extracción | 22 |
| 3.2.2. Diseño experimental planteado para la optimización de la extracción | 23 |
| 3.3. Cuantificación y pureza del ADN extraído | 25 |
| 3.4. Amplificación del gen <i>CSN3</i> mediante PCR en tiempo real (qPCR) | 25 |
| 3.4.1. Diseño de cebadores y sondas | 26 |
| 3.4.2. Diseño de muestras de referencia comerciales de ADN sintético | 28 |
| 3.4.3. Optimización de las condiciones de amplificación mediante qPCR | 28 |
| 3.5. Validación del método de identificación de alelos A, B, y E del gen <i>CSN3</i> mediante qPCR | 33 |
| 3.5.1. Rango dinámico, límite de detección (LOD) y cuantificación (LOQ) | 33 |
| 3.5.2. Eficiencia de la amplificación | 33 |
| 3.5.3. Repetibilidad, intra-assay e inter-assay repeatability | 34 |
| 3.6. Aplicabilidad de la qPCR en la identificación de los alelos A, B y E de la kappa-caseína | 34 |
| 4. RESULTADOS | 35 |
| 4.1. Extracción de ADN a partir de células somáticas de leche y de pelo | 35 |
| 4.2. Optimización de las condiciones de amplificación mediante qPCR | 37 |
| 4.2.1. Optimización de la concentración de cebadores | 37 |
| 4.2.2. Optimización del funcionamiento de las sondas | 38 |
| 4.3. Validación del método de identificación de alelos A, B y E mediante qPCR | 40 |
| 4.4. Aplicabilidad de la qPCR en la identificación de los alelos A, B y E de la kappa-caseína | 43 |
| 5. CONCLUSIONES | 45 |
| BIBLIOGRAFÍA | 47 |
| ANEXOS | 51 |

ÍNDICE DE TABLAS

| | |
|--|----|
| Tabla 1. Protocolo de extracción del kit NucleoSpin Tissue..... | 23 |
| Tabla 2. Matriz del diseño experimental planteado para la optimización de la extracción..... | 24 |
| Tabla 3. Secuencia del amplicón, primers y sondas específicas; diseñados para la puesta a punto de la qPCR para la identificación de las variantes alélicas A, B y E de la kappa-caseína. | 27 |
| Tabla 4. Características del amplicón, primers y sondas específicas, diseñados para la puesta a punto de la qPCR para la identificación de las variantes alélicas A, B y E de la kappa-caseína. | 27 |
| Tabla 5. Estructuras secundarias detectadas con la herramienta “OligoAnalyzer” en el diseño de la qPCR para la identificación de las variantes alélicas A, B y E de la kappa-caseína. . | 27 |
| Tabla 6. Muestras de referencia sintéticas de las variantes A, B y E de la kappa-caseína empleadas como controles..... | 28 |
| Tabla 7. Matriz de cebadores para la optimización de la amplificación del gen <i>CSN3</i> | 29 |
| Tabla 8. Volumen de los reactivos (µl) para la optimización de la concentración de los cebadores para la amplificación del gen <i>CSN3</i> mediante PCR en tiempo real. | 29 |
| Tabla 9. Conjunto de pruebas realizadas para la optimización de las cuatro sondas: FAM_13104_A, HEX_13104_C, Cy5_13124_A y FAM_13124_G..... | 31 |
| Tabla 10. Volumen de los reactivos (µl) para la optimización del funcionamiento de las sondas en reacciones en singleplex. | 31 |
| Tabla 11. Volumen de los reactivos (µl) para la optimización del funcionamiento de las sondas en reacciones en duplex..... | 32 |
| Tabla 12. Análisis de la varianza del factor “extracción”. | 35 |
| Tabla 13. Efecto de la temperatura de conservación de las muestras y tipo de matriz en la cantidad (ng/µl), calidad ($A_{260/280}$ y $A_{260/230}$) y <i>amplificability</i> (Cq) de las muestras de ADN. | 35 |
| Tabla 14. Análisis de la varianza del factor “extracción”. | 37 |
| Tabla 15. Efecto de la concentración de cebadores y cantidad de muestra en valores Cq y niveles de fluorescencia alcanzados al final de la amplificación (End RFU) de las muestras de ADN. | 37 |
| Tabla 16. Resultados de la prueba N°7 en las que se fijan las condiciones de las cuatro sondas para la detección de las variantes A, B y E de la kappa-caseína. | 39 |
| Tabla 17. Intervalos del Coeficiente de Variación (%) de la repetibilidad intra- e inter-assay, entre los que se encuentran los valores del rango dinámico fijado para los tres alelos y la pareja de sondas de la posición 13104 (FAM_13104_A y HEX_13104_C)..... | 42 |
| Tabla 18. Intervalos del Coeficiente de Variación (%) de la repetibilidad intra- e inter-assay, entre los que se encuentran los valores del rango dinámico fijado para los tres alelos y la pareja de sondas de la posición 13124 (Cy5_13124_A y FAM_13124_G)..... | 43 |
| Tabla 19. Señal esperada y valores medios Cq obtenidos en las curvas de amplificación, y alelos esperados en los gráficos de discriminación alélica; para los seis posibles genotipos del gen <i>CSN3</i> | 43 |
| Tabla 20. Resultados individuales para cada sonda en la optimización de su concentración y condiciones de tiempo y temperatura de la fase annealing. | 51 |
| Tabla 21. ANOVA simple del factor “concentración”, de las sondas HEX_13104_C y Cy5_13124_A. | 51 |
| Tabla 22. Comparativa de los valores End RFU para la selección de la concentración estadísticamente mejor. | 51 |
| Tabla 23. Pendiente, Eficiencia (%) y valor de R^2 de los tres ensayos realizados para las sondas de la posición 13104 (FAM_13104_A y HEX_13104_C). Valores de LOD, LOQ y rangos dinámicos fijados para cada alelo y esta pareja de sondas..... | 52 |

| | |
|---|----|
| Tabla 24. Pendiente, Eficiencia (%) y valor de R^2 de los tres ensayos realizados para las sondas de la posición 13124 (Cy5_13124_A y FAM_13124_G). Valores de LOD, LOQ y rangos dinámicos fijados para cada alelo y esta pareja de sondas..... | 52 |
| Tabla 25. Repetibilidad intra-assay de las curvas estándar de las sondas de la posición 13104 (FAM_13104_A y HEX_13104_C), para validar la discriminación alélica de las variantes A, B y E de la kappa-caseína. ** | 53 |
| Tabla 26. Repetibilidad intra-assay de las curvas estándar de las sondas de la posición 13124 (Cy5_13124_A y FAM_13124_G), para validar la discriminación alélica de las variantes A, B y E de la kappa-caseína. ** | 53 |
| Tabla 27. Repetibilidad inter-assay de las curvas estándar de las sondas de la posición 13104 (FAM_13104_A y HEX_13104_C), para validar la discriminación alélica de las variantes A, B y E de la kappa-caseína. ** | 54 |
| Tabla 28. Repetibilidad inter-assay de las curvas estándar de las sondas de la posición 13124 (Cy5_13124_A y FAM_13124_G), para validar la discriminación alélica de las variantes A, B y E de la kappa-caseína. ** | 54 |
| Tabla 29. ANOVA simple para evaluar el efecto del tipo de matriz en la cantidad (ng/μl) y ($A_{260/280}$) calidad de ADN extraído. | 55 |
| Tabla 30. Cantidad (ng/μl) y calidad de ADN extraído ($A_{260/280}$) de las muestras de leche y pelo procedentes de los mismos individuos genotipados. | 55 |

ÍNDICE DE FIGURAS

| | |
|---|----|
| Figura 1. Organización estructural de las unidades de transcripción que codifican las 6 principales proteínas de la leche..... | 8 |
| Figura 2. Posición dentro del gen y de la proteína de las variantes genéticas del <i>CSN3</i> para especies del género <i>Bos</i> | 10 |
| Figura 3. Comparación de las secuencias nucleotídicas entre las variantes alélicas A y B, y A y E; del gen <i>CSN3</i> en la especie <i>Bos Taurus</i> | 11 |
| Figura 4. Curva de disociación de una qPCR realizada en laboratorio. | 18 |
| Figura 5. Mecanismo de discriminación alélica mediante el empleo de sondas TaqMan (BIO-RAD Laboratories, 2006). | 19 |
| Figura 6. Gráfico de amplificación de una qPCR realizada en laboratorio. | 19 |
| Figura 7. Programa de amplificación (Temperatura-Tiempo) para la amplificación del gen <i>CSN3</i> con el agente SYBR Green..... | 30 |
| Figura 8. Curva de disociación de la qPCR. | 37 |
| Figura 9. Curvas de amplificación (RFU frente al número de ciclos) de las curvas estándar y gráfico de discriminación alélica (RFU del fluoróforo HEX frente al RFU del fluoróforo FAM) del segundo ensayo..... | 40 |
| Figura 10. Curvas de amplificación (RFU frente al número de ciclos) de las curvas estándar y gráfico de discriminación alélica (RFU del fluoróforo Cy5 frente al RFU del fluoróforo FAM) del tercer ensayo. | 41 |
| Figura 11. Curvas estándar, con el rango dinámico ajustado, de los tres ensayos realizados para la pareja de sondas de la posición 13104: FAM_13104_A y HEX_13104_C (izquierda) y la pareja de sondas de la posición 13124: Cy5_13124_A y FAM_13124_G (derecha). | 42 |
| Figura 12. Gráficos de discriminación alélica de las parejas de sondas de la posición 13104 y 13124 (arriba RFU de la sonda HEX_13104_C frente al RFU de la sonda FAM_13104_A, debajo RFU de la sonda Cy5_13124_A frente al RFU de la sonda FAM_13124_G). | 44 |

ABSTRACT

Kappa-casein is a milk protein encoded by the *CSN3* gene, for which 14 polymorphisms have been described. Specifically, in the Friesian breed, the variants of greatest interest due to their frequency and influence on the technological aptitude of milk are the A, B and E alleles. Therefore, especially in farms that process dairy products, it is of interest genetic selection focused on increasing the frequency of the B allele, for which favourable effects on coagulation properties and cheese yield are described, and decreasing the A and E alleles, associated with negative effects. In the present work, the qPCR technique using probes for the detection of the A, B and E alleles of the *CSN3* gene in the Friesian breed has been developed. Primers and specific TaqMan probes were designed to identify the three variants in duplex reactions. Under the final concentrations of 0,2 μM of primers and 0,4 μM of probe and after adjusting the thermo-cycling protocol, the detection limit was set at $6,73 \times 10^0$ copies/ μl , the defined dynamic ranges showed efficiency values close to 100% and CV (%) values between 0,01 and 15,76, showing a good repeatability of the method. The applicability of the technique showed good results in the discrimination of variants using the 13104 probes, but the suboptimal performance of the 13124 probes led to imprecise identification of some genotypes. Finally, the possibility of redesigning the FAM_13124_G probe was suggested for further refinement of the technique in future assays.

Keywords: milk proteins, kappa-casein, *CSN3* gene, genotyping, real-time PCR, Holstein-Friesian.

RESUMEN

La kappa-caseína es una proteína láctea codificada por el gen *CSN3*, para el que se han descrito 14 polimorfismos. Concretamente, en la raza Frisona, las variantes de mayor interés debido a su frecuencia e influencia en las aptitudes tecnológicas de la leche son los alelos A, B y E. A tal efecto, especialmente en granjas que transforman derivados lácteos, es de interés la selección genética enfocada en el incremento de la frecuencia del alelo B, para el que se describen efectos favorables sobre las propiedades de coagulación y rendimiento quesero, y la disminución de los alelos A y E, asociados a efectos negativos. En el presente trabajo se ha realizado la puesta a punto de la técnica qPCR mediante sondas para la detección de los alelos A, B y E del gen *CSN3* en la raza Frisona. Se diseñaron cebadores y sondas específicas TaqMan que permitiesen identificar las tres variantes en reacciones duplex. Bajo las concentraciones finales de 0,2 μM de cebadores y 0,4 μM de sonda, y tras ajustar los programas del termociclador; el límite de detección se fijó a $6,73 \times 10^0$ copias/ μl , los rangos dinámicos definidos manifestaron valores de eficiencia cercanos al 100% y valores de CV (%) comprendidos entre 0,01 y 15,76, observando una buena repetibilidad del método. La aplicabilidad de la técnica reflejó buenos resultados en la discriminación de las variantes mediante las sondas de la posición 13104, pero el funcionamiento no tan óptimo de las sondas 13124, supuso la identificación imprecisa de algunos genotipos. Finalmente, se sugirió la posibilidad de rediseñar la sonda FAM_13124_G para lograr el perfeccionamiento de la técnica en futuros ensayos.

Palabras clave: proteínas lácteas, kappa-caseína, gen *CSN3*, genotipado, PCR en tiempo real Holstein-Friesian.

1. INTRODUCCIÓN

1.1. La leche y los derivados lácteos

La leche es un sistema polifásico compuesto por agua (87%), grasa emulsionada (3,68%), micelas proteicas de caseína en estado coloidal y proteínas en solución (3,51%), lactosa (4,98%), y sales y micronutrientes (0,47%). Debido a su elevado contenido en lípidos, proteínas de calidad, vitaminas liposolubles y sales minerales, la leche es una fuente importante de macro y micronutrientes, algo muy importante en la alimentación de los mamíferos neonatos para su correcto desarrollo y crecimiento (Giglioti et al., 2020; Sah et al., 2018; Sebastiani et al., 2020). No obstante, el consumo de este producto es interesante para cualquier tipo de consumidor: aporta calcio, vitamina D, proteínas que favorecen la mineralización ósea, su consumo se asocia a un menor riesgo de obesidad y presenta beneficios sobre el riesgo cardiometabólico (Rojo Martínez, 2019). En definitiva, dadas sus características nutricionales y los beneficios a la salud que se le atribuyen, la leche de vaca constituye un producto fundamental en la alimentación humana.

La demanda y consumo de leche continúa aumentando, de tal forma que durante la próxima década, se prevé un crecimiento anual de la producción mundial en un 1,7% (Organisation for Economic Co-operation Development (OECD) and Food and Agricultural Organization of the United Nations (FAO), 2021, Chapter 7). No obstante, no toda la leche se destina a su consumo directo, sino también a la elaboración de derivados lácteos como yogures, queso o mantequilla (Miranda et al., 2020). Según los datos reportados por el Ministerio de Agricultura Pesca y Alimentación (MAPAMA), el consumo per cápita de productos lácteos en España durante el año 2020 fue 37,41 kg/persona. Aunque la cuota de consumo fue inferior a la leche líquida, la proporción en valor de la facturación total de leche y derivados lácteos fue de 138,22 € en derivados lácteos, frente a 51,12 € de leche líquida. Por otra parte, entre los diferentes derivados, el consumo per cápita de queso representó una proporción muy importante (8,80 kilos/persona), suponiendo la evolución favorable de su volumen y valor, respecto al año anterior (un incremento del 13% y 15%, respectivamente) (MAPAMA, 2021). Según reportan otros datos más recientes, en el mes de mayo de 2021, solo el 44% de leche se destinó a consumo directo y el porcentaje restante a derivados lácteos. Entre ellos, el queso con un 18%, fue el derivado de mayor porcentaje (Subdirección General de Producciones Ganaderas y Cinegéticas & Dirección General de Producciones y Mercados Agrarios, 2021). Estos datos no solo son aplicables a nivel nacional, sino también a escala mundial. Según recogen los últimos datos sobre las perspectivas agrícolas para los años 2021-2030, el queso constituye el segundo producto lácteo más importante, producido principalmente en Europa y Norteamérica; y de la misma forma, se espera que su consumo per cápita continúe aumentando (OECD & FAO, 2021, Chapter 7).

Una de las razones por las que el consumo de los derivados lácteos ha sido tan extendido a lo largo de la historia, es la elevada perecibilidad de la leche. Esto obliga a su inmediato procesamiento o conservación una vez recogida, lo cual ha supuesto que los derivados hayan sido empleados como buenos métodos de conservación. Al fin y al cabo, la fermentación es una de las técnicas más antiguas de conservación de alimentos, la cual ha supuesto a su vez, el origen de una infinita variedad de productos lácteos. Como se ha mencionado, se ha comprobado que el queso es uno de los derivados más importantes en cuanto a producción y consumo, y que en comparación con la leche líquida, proporciona otros beneficios nutricionales adicionales como una mejor digestibilidad de la lactosa, proteínas, lípidos o de algunas vitaminas que lo componen. Además, estos también pueden ser conductores de probióticos, los cuales se asocian al antagonismo de microorganismos patógenos, obtención de energía de compuestos no digeribles por los humanos, producción de metabolitos beneficiosos como ácidos orgánicos o vitaminas, y modulación del

sistema inmune (Alonso López & Mayo Pérez, 2015). En definitiva, la vida útil y beneficios nutricionales que ofrecen productos como el queso, han contribuido a que sea ampliamente consumido y a que su demanda continúe aumentando.

1.2. Las proteínas lácteas

Las proteínas lácteas son un compuesto de gran interés, debido a su papel fundamental en las propiedades nutricionales y tecnológicas de la leche. Desde el punto de vista nutricional, estas presentan un elevado valor biológico ya que contienen todos los aminoácidos esenciales (Vougiouklaki et al., 2020). Por otra parte, también existen evidencias acerca de moléculas de naturaleza proteica y especialmente caseínica, que logran convertir la leche en un vector importante de compuestos bioactivos. Estos pueden ser proteínas intactas, o péptidos que se encuentran encriptados en las cadenas aminoacídicas de las proteínas y que son liberados a través de su digestión enzimática en el tracto gastrointestinal, a través de la actividad proteolítica de los microorganismos empleados para la fermentación, o a través de enzimas empleadas para el procesamiento de alimentos (como por ejemplo la quimosina). El interés de dichos compuestos radica en los beneficios a los que se le asocia, participan en diferentes funciones fisiológicas, presentan propiedades antioxidantes y antimicrobianas, además de efectos inmunoestimuladores que reducen el riesgo de desarrollar ciertas enfermedades (efecto de interés en los recién nacidos ya que su sistema inmunitario y capacidad de defensa permanecen inmaduros tras el nacimiento) (Korhonen & Pihlanto, 2006; Miranda et al., 2020; Sah et al., 2018). Con respecto a las propiedades tecnológicas, el contenido proteico también es un factor clave en el procesamiento industrial de los derivados, ya que condiciona directamente las aptitudes de transformación de la leche. En este ámbito, es de nuevo destacable la función de las caseínas, ya que condicionan la formación del cuajo durante la elaboración del queso (Miranda et al., 2020).

Usualmente, las caseínas han sido relacionadas con el suministro de aminoácidos, fosfato y calcio, dada su composición; y se le ha atribuido la función de retener los ácidos grasos de la leche mediante la formación de micelas (Cortés-López et al., 2012; Miranda et al., 2020). No obstante, como se ha comprobado, las evidencias acerca de sus interesantes propiedades tecnológicas y las propiedades antioxidantes y antimicrobianas de sus hidrolizados, continúan aumentando. Por este motivo, las caseínas han constituido y constituyen un interesante objeto de estudio.

Como se ha mencionado, el contenido proteico de la leche es elevado. Este abarca más de 400 compuestos diferentes, clasificables en tres grupos: las caseínas (80-85%), organizadas como estructuras supramoleculares dinámicas que encierran fosfato de calcio coloidal, denominadas micelas; las proteínas séricas (13-18%), disueltas en la fase acuosa; y las proteínas que constituyen parte de la bicapa lipídica de los glóbulos grasos de leche (1-2%), protegiendo su membrana de efectos de coalescencia y lipólisis (Bär et al., 2019). Las caseínas se subdividen a su vez en cuatro grupos o fracciones, α_1 -CN, α_2 -CN, β -CN, y κ -CN; codificados respectivamente por los genes *CSN1S1*, *CSN1S2*, *CSN2* y *CSN3* (Adamov et al., 2020; Sebastiani et al., 2020). Este es un grupo de genes de 250 kb, localizado en el cromosoma 6, denominado locus CN o super locus (Caroli et al., 2009). En cuanto a las proteínas del suero, existen dos compuestos principales, la α -lactalbúmina (α -LA) y la β -lactoglobulina (β -LG): codificados respectivamente por los genes *LAA*, situado en el cromosoma 5, y *LGB*, situado en el cromosoma 11 (Caroli et al., 2009). De esta forma, la fracción mayoritaria de proteínas sintetizadas por las células epiteliales mamarias, está codificada en los seis genes estructurales representados en la Figura 1.

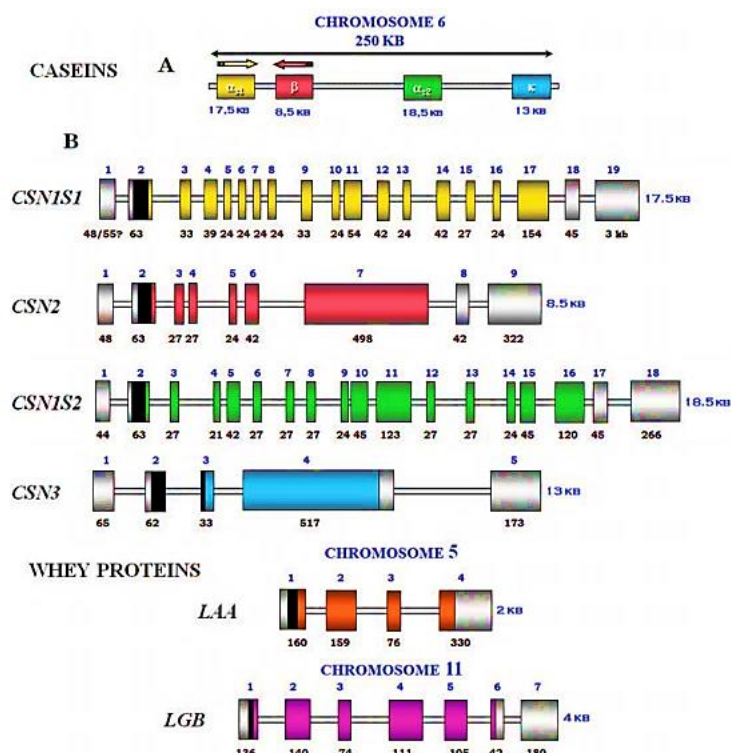


Figura 1. Organización estructural de las unidades de transcripción que codifican las 6 principales proteínas de la leche. A) Organización genómica del locus de la caseína bovina. B) Organización estructural de las 6 unidades transcripcionales de la proteína láctea. Las barras abiertas representan los intrones; los exones se representan con recuadros grandes, grises (regiones no traducidas 5' y 3'), negros (parte del exón que codifica el péptido señal) y coloreados (exones y parte de los exones que codifican las proteínas maduras). El tamaño de los exones se indica, en pares de bases, debajo de cada exón con su número indicado en la parte superior (Caroli et al., 2009).

Según la bibliografía publicada, las caseínas son muy polimórficas en el ganado vacuno. Esto quiere decir, que existe un gran número de variantes genéticas de las caseínas en la leche bovina (Miranda et al., 2020). Dichas variaciones genéticas o polimorfismos, surgen como resultado de la inserción, delección, duplicación o inversión de fragmentos de ADN de diferentes longitudes; o de la sustitución de un solo nucleótido en la secuencia de ADN. El cambio de un nucleótido son polimorfismos denominados SNP (single nucleotide polymorphism), y dependiendo de si afectan o no a la secuencia aminoacídica, son respectivamente mutaciones no sinónimas o sinónimas (Food and Agriculture Organization of the United Nations (FAO), 2007, Part 4). Se ha demostrado que la interacción entre las variaciones dentro de las secuencias codificantes y no codificantes del genoma, regulan la expresión génica global de la proteína específica, y consecuentemente, influyen en la composición nutricional de la leche y en las propiedades tecnológicas que determinan los procesos de elaboración del queso. Por este motivo, debido a su gran variabilidad y a su asociación directa con los rasgos productivos y nutricionales de la leche, numerosas investigaciones se han enfocado en la descripción del locus CN y en el estudio de su variabilidad genética, teniendo en cuenta el conjunto del haplotipo (Caroli et al., 2009).

Finalmente, se han descrito un gran número de polimorfismos con importantes diferencias entre especies y razas del género *Bos*. Por consiguiente, es preciso denotar que no todas las variantes hasta ahora descubiertas, se encuentran presentes en todas las especies o razas estudiadas hasta el momento: *Bos taurus* (bovino), *Bos indicus* (cebú), *Bos grunniens* (yack) y *Bos javanicus* (banteng de bali). En cualquier caso, actualmente se han identificado 8 variantes de la proteína

α_{s1} -CN (A, B, C, D, E, F, G, H), 4 de la α_{s2} -CN (A, B, C, D), 12 de la β -CN (A^1 , A^2 , A^3 , B, C, D, E, F, G, H^1 , H^2 , I), 11 de la κ -CN (A, B, C, E, F^1 , F^2 , G^1 , G^2 , H, I, J), 11 de la β -LG (A, B, C, D, E, F, G, H, I, J, W), y 3 de la α -LA (A, B, C) (Caroli et al., 2009).

Puesto que dichos polimorfismos influyen en las características de la leche, consecuentemente también lo hacen en aspectos económicos relativos a la producción de sus derivados. En lo que respecta a sus aptitudes en la elaboración del queso, la leche debe reunir un conjunto de propiedades tecnológicas, físicas y químicas e higiénicas. Estas permiten la elaboración del queso o requesón, mediante la formación de un coágulo denso de caseína por medio de la adición de una enzima de cuajo (Tyulkin et al., 2018). De esta forma, muchos trabajos se han centrado en el estudio de la calidad reológica de la leche, asociable a las diferentes variantes proteicas de las caseínas. Concretamente, estos se han enfocado en las kappa-caseínas (κ -CN), puesto que se encuentran predominantemente en la superficie de las micelas caseínicas y son el sustrato específico de la quimosina. Esta es una enzima de cuajo cuya actividad hidrolítica divide su cadena aminoacídica en la posición Phe105-Met106, dando lugar a una fracción insoluble (para- κ -CN: aminoácido 1-105) y a un caseinomacropéptido-CMP soluble (CMP: aminoácido 106-169). Esto no solo es crucial en el proceso de producción del queso, sino que también denota cierto interés nutricional, ya que a la fracción CMP se le atribuyen funciones fisiológicas, como una mayor digestibilidad del queso, además de propiedades antibacterianas (Adamov et al., 2020; Caroli et al., 2009). Por otra parte, estas proteínas también influyen en la resistencia de la leche frente a elevadas temperaturas. Esto es indispensable, por ejemplo, en el procesado de aquellos productos en los que se pretende prolongar su vida útil, ya que impiden que se produzca la coagulación proteica durante su esterilización (Tyulkin et al., 2018).

Tal y como reflejan las tendencias de consumo, desde hace algunos años los consumidores demandan productos de mayor calidad. Como se ha podido comprobar, en el sector lácteo la calidad de la leche y sus derivados, está directamente condicionada por un mayor contenido proteico, especialmente caseínico y más concretamente de la kappa-caseína (Requena et al., 2007). De esta forma, en la ganadería e industria láctea, para lograr satisfacer la exigente demanda de los actuales consumidores y lograr una posición competitiva en el mercado, es de interés la identificación de aquellas variantes del gen de la kappa-caseína o el gen *CSN3* que logran optimizar las propiedades reológicas de la leche.

1.3. La Kappa-caseína

La kappa-caseína es una fosfoproteína conformada por 169 aminoácidos, constituye entorno al 12% de la fracción caseínica y se encuentra codificada por el gen *CSN3* localizado en el cromosoma 6 (Figura 1). Dicho gen presenta una longitud de 13 kb y está dividido en una unidad de transcripción (5 exones y 4 intrones) y las regiones no transcripcionales 5' y 3'. El cuarto exón, con una longitud de 517 pb, alberga las 11 mutaciones no sinónimas o sustituciones de un solo nucleótido, que hasta la actualidad han podido identificarse en el género *Bos*. De esta forma, existen 11 variantes diferentes de la kappa-caseína: A, B, C, E, F^1 , F^2 , G^1 , G^2 , H, I y J. No obstante, es preciso denotar que en una última revisión, se señaló la existencia de otros dos alelos resultantes de mutaciones no sinónimas, los alelos B^2 y D. También se ha podido identificar una sustitución nucleotídica sinónima o silenciosa, la cual no implica cambios en la secuencia aminoacídica, que corresponde al alelo A^1 (Adamov et al., 2020). Todas las mutaciones mencionadas, así como su localización en el gen y proteína, son representadas en la Figura 2 tomando como referencia la secuencia nucleotídica o aminoacídica del alelo A.

| Gene, Protein | <i>CSN3</i> variant ² | | | | | | | | | | | | | |
|---------------|----------------------------------|-----|-----|-----|-----|-----|-----|----------------|----------------|----------------|----------------|-----|-----|-----|
| | A | A' | B | B' | C | D | E | F ^d | F ^e | G ^f | G ^g | H | I | J |
| 12690 | CGC | | | | | | | | CAC | | | | | |
| 10 | Arg | | | | | | | | His | | | | | |
| 12940 | ACT | | | ACC | | | | | | | | | | |
| 93 | Thr | | | Thr | | | | | | | | | | |
| 12950 | CGT | | | | | | | | | TGT | | | | |
| 97 | Arg | | | | | | | | | Cys | | | | |
| 12951 | CGT | | | | CAT | CAT | | | | | | | | |
| 97 | Arg | | | | His | His | | | | | | | | |
| 12971 | TCA | | | | | | | | | | | | GCA | |
| 104 | Ser | | | | | | | | | | | | Ala | |
| 13065 | ACC | | | | | | | | | ATC | | ATC | | |
| 135 | Thr | | | | | | | | | Ile | | Ile | | |
| 13068 | ACC | | ATC | ATC | ATC | | | | | | | | | ATC |
| 136 | Thr | | Ile | Ile | Ile | | | | | | | | | Ile |
| 13096 | ACT | | | | | | | ACG | | | | | | |
| 145 | Thr | | | | | | | | | | | | | |
| 13104 | GAT | | GCT | GCT | GCT | | | GTT | | | GCT | | | GCT |
| 148 | Asp | | Ala | Ala | Ala | | | Val | | | Ala | | | Ala |
| 13111 | CCA | CCG | | | | | | | | | | | | |
| 150 | Pro | | | | | | | | | | | | | |
| 13119 | ATT | | | ACT | | | | | | | | | | |
| 153 | Ile | | | Thr | | | | | | | | | | |
| 13124 | AGC | | | | | | | | | | | | | |
| 155 | Ser | | | | | | GGC | | | | | | | |
| 13162 | ACT | | | ACC | | | Gly | | | | ACC | | | Arg |
| 167 | Thr | | | | | | | | | | | | | |
| 13165 | GCA | | GCG | GCG | GCG | ? | | | | | GCG | | | |
| 168 | Ala | | | | | | | | | | | | | |

Figura 2. Posición dentro del gen (tomando como referencia la numeración desde el inicio de la región codificante o CDS) y de la proteína de las variantes genéticas del *CSN3* para especies del género *Bos*. La secuencia empleada como referencia es el alelo A. Las mutaciones no idénticas e idénticas se encuentran representadas en negrita y cursiva, respectivamente, para cada codón (Caroli et al., 2009).

Según sugiere la bibliografía publicada, los efectos de los polimorfismos proteicos en las propiedades físico-químicas y queseras de la leche, pueden ser de dos tipos: un efecto biológico directo de la variante proteica, por cambio en las características físico-químicas de la leche; o un efecto cuantitativo directo de la mutación del gen, por variación en la tasa de síntesis de la proteína. En el caso de los polimorfismos de la kappa-caseína, su efecto se ajusta al segundo tipo. Se ha demostrado que hay variantes que con respecto a otras, presentan un mayor nivel de expresión en la glándula mamaria (Caroli et al., 2009; Medrano, 2000). Concretamente, según han reportado diferentes publicaciones, las dos variantes más comunes de la kappa-caseína, las variantes A y B, son las que generan un mayor efecto sobre las propiedades queseras de la leche y su estabilidad frente al calor. Esto indica que uno de los dos alelos supone una mayor síntesis de su variante proteica frente al otro, y por ello un efecto diferente en las propiedades tecnológicas, ya que la cantidad total de kappa-caseínas determina la estructura y tamaño de las micelas. Aparentemente, una mayor cantidad de kappa-caseínas favorece a la formación de micelas más pequeñas y uniformes, lo cual contribuye a un queso más termoestable y a la formación de un coágulo firme y denso que retiene más sólidos e incrementa el rendimiento en la producción de quesos (Caroli et al., 2009; Medrano, 2000; Requena et al., 2007).

Por otra parte, se ha evidenciado que alelos menos comunes, también podrían afectar a las propiedades reológicas de la leche. Este es el caso de algunas variantes identificadas en la raza Frisona (Holstein-Friesian o HF), para el que se definieron efectos negativos (Caroli et al., 2009). La raza Frisona es la raza de bovino de leche más extendida a nivel mundial, dado que por medio de sucesivos programas de selección, se logró incrementar considerablemente su rendimiento productivo (FAO, 2007, Part 1). Por este motivo, la mayoría de los estudios publicados sobre la especie *Bos taurus*, emplean especímenes de esta raza. En numerosas poblaciones de la raza Frisona (macedonia, china, iraní, eslovaca, canadiense, croata, egipcia, argentina, rumana, serbia, indonesia, rusa, polaca, italiana y francesa), las variantes A y B son también las variantes más comunes (Adamov et al., 2020; Chessa et al., 2020; Sanchez et al, 2020). Por el contrario, el alelo E, es la variante que presenta una menor frecuencia y al que se le han asociado efectos negativos (Adamov et al., 2020; Caroli et al., 2009). Pese a que existen pequeñas diferencias entre las poblaciones de esta raza (0,546 (alelo A), 0,317 (alelo B) y 0,131 (alelo E) en la población italiana;

0,569 (alelo A), 0,336 (alelo B) y 0,095 (alelo E) en la población francesa; o 0,584 (alelo A), 0,336 (alelo B) y 0,08 (alelo E) en la población macedonia (Adamov et al., 2020; Chessa et al., 2020; Sanchez et al., 2020)), las frecuencias de los tres alelos siempre guardan proporciones similares. Concretamente, las frecuencias alélicas estimadas para las poblaciones de la raza Frisona a nivel mundial son 0,520 para el alelo A, 0,335 para el alelo B y 0,139 para el alelo E (Chessa et al., 2020). En definitiva, se puede considerar que las variantes de mayor interés debido a su frecuencia y/o influencia en las aptitudes tecnológicas de la leche, son los alelos A, B y E.

Tal y como representa la Figura 3, las variantes A, B y E del gen *CSN3*, son resultado de mutaciones no sinónimas. Dichas mutaciones consisten en la sustitución de un nucleótido de la secuencia genómica, lo cual conlleva al cambio del codón y el consecuente cambio del aminoácido de la cadena proteica. De esta forma, la variante proteica B difiere de la variante A en dos posiciones aminoacídicas, como consecuencia de dos mutaciones. En la posición 136 de la cadena aminoacídica de la variante B, en lugar de una Threonina (Thr) hay una Ileacina (Ile). Esto se debe a la sustitución del nucleótido Citosina de la posición 13068, por una Tiamina, lo cual conlleva el cambio del codon ACC que codifica la Thr, al codon ATC que codifica la Ile. Por otra parte, en la posición aminoacídica 148, la Asparagina (Asp) se intercambia por la Alanina (Ala), dado que el nucleótido Adenina es intercambiado por una Citosina, modificando el codon de GAT a GCT (posición 13104). En cuanto a la variante E, esta difiere del alelo A en sólo un aminoácido, el aminoácido 155. En este caso, la sustitución de una Adenina por una Guanina (posición 13124) supone el cambio de los codones AGC a GGC, lo cual conlleva a que en la secuencia aminoacídica la Serina (Ser) sea intercambiada por la Glicosina (Gly) (Adamov et al., 2020). En la Figura 3 se representan estas mutaciones no sinónimas comparando las secuencias nucleotídicas de los alelos B y E, con respecto al alelo A.



Figura 3. Comparación de las secuencias nucleotídicas entre las variantes alélicas A y B, y A y E; del gen *CSN3* en la especie *Bos Taurus*. La posición se numera desde el inicio de la CDS. Secuencias extraídas de la base de datos “GenBank” ((National Center for Biotechnology Information (NCBI), 2021b)).

1.4. Evolución del proceso de selección genética del ganado bovino y las propiedades reológicas de la leche

En la industria láctea la mejora de la producción de leche y de su composición es un objetivo fundamental, y por ello, un importante criterio en la selección animal. En este sentido, en los países desarrollados, los primeros programas de mejora surgieron a inicios del siglo XIX (FAO, 2007, Part 1). Inicialmente, la selección basaba sus predicciones en la información fenotípica, datos genealógicos y herramientas estadísticas (González-Recio et al., 2009; Medrano, 2000). De esta forma, la selección de toros y vacas lecheras, se realizaba en base a criterios cuantitativos tales como los rendimientos de leche, proteína y grasa; seleccionando para el cruzamiento aquellos animales que generasen los mejores rendimientos productivos. Posteriormente, gracias a los avances tecnológicos, también se comenzó a comercializar el semen y embriones de dichos animales. No obstante, la selección basada exclusivamente en las propiedades de los productos

obtenidos, suponía una mejora genética relativamente lenta (FAO, 2007, Part 1; Volkandari et al., 2017).

A finales del siglo XX, a raíz de los revolucionarios avances en el campo de la biología molecular, se comenzó a plantear que el estudio desde una perspectiva molecular podría lograr una selección más eficiente y, consecuentemente, una mejora genética más rápida. Conocer los genes (locus, estructura, transmisión) y sus genotipos, permitiría analizar la variabilidad genética y parámetros genéticos, como la heredabilidad y la correlación con otros caracteres cuantitativos. A esto se sumó el descubrimiento de los marcadores moleculares, un locus que presenta polimorfismos y se hereda de manera mendeliana. Todo ello permitió la realización de investigaciones para mapear e identificar genes que afectasen a características cuantitativas de importancia económica (Abdel-Azim & Freeman, 2002; Al-Khudhair, et al., 2021). De esta manera, si posteriormente se confirmaba la correlación de dichos polimorfismos con una característica productiva, podrían emplearse como marcadores moleculares en los que basar la selección y cría del ganado (González-Recio et al., 2009; Medrano, 2000).

Sendas investigaciones lograron identificar locis para características o rasgos cuantificables, que permitió considerarlos como marcadores de interés económico y que fueron denominados QTL (Quantitative Trait Loci). Los QTL se pueden utilizar para llevar a cabo la genotipificación de individuos en las poblaciones de especies animales, entre los que se puede detectar a aquellos que son portadores de estos marcadores para ser seleccionados y con ello desarrollar programas de selección. A este tipo de selección se le ha denominado selección asistida por marcadores (SAM) o selección asistida por genes (SAG) (Miluchová et al., 2018; Requena et al., 2007; Volkandari et al., 2017). Sin embargo, su aplicación para características complejas de determinación poligénica ha sido limitada, por la pequeña proporción de la variación fenotípica que explican cada uno de los marcadores (Goddard & Hayes, 2007), dado el escaso número de marcadores utilizados tradicionalmente. Así, a principios de este siglo, nace la selección genómica (Meuwissen et al., 2001), como una alternativa para aumentar la eficacia de la SAM. Los métodos de evaluación genómica, además de la información fenotípica y de pedigrí, utilizan la información genotípica obtenida mediante chips de alta densidad (desde 3.000 hasta 700.000 SNPs), que recopilan información sobre miles de marcadores SNPs y que están relacionados con el mérito genético de los animales. Asimismo, en los citados chips incluyen también información de genotipos de caracteres de interés como las proteínas lácteas o enfermedades genéticas (Boichard et al., 2018).

En el caso de las investigaciones relativas a las proteínas lácteas, el cromosoma 6 logró identificarse como un QTL para aspectos relativos a cantidad de leche producida y a los porcentajes de contenido graso y proteico (Fontanesi et al., 2014; Spelman et al., 1996). Tal y como se mencionaba con anterioridad, estos primeros estudios consideraron el haplotipo completo del locus CN, en lugar de estudiar individualmente los genes que codifican las cuatro caseínas (Caroli et al., 2009). De esta forma, las investigaciones posteriores continuaron profundizando mediante el estudio de los locus individuales (Miluchová et al., 2018; Volkandari et al., 2017). Sin embargo, es preciso señalar que a pesar de ser extensamente estudiada, la información genética de las proteínas de la leche no se ha tenido muy en cuenta en los programas de selección del vacuno lechero. No ha sido hasta hace relativamente unos años, que la industria ha comenzado a ganar interés en la mejora de la calidad quesera mediante la selección genética en algunos países, como Italia (Chessa et al., 2020; Cole et al., 2020; Hayes et al., 2009).

En este ámbito, actualmente existe una amplia bibliografía de estudios publicados que se fundamentan en la determinación de los rasgos productivos de la leche, producida por diferentes

poblaciones de diferentes razas de ganado bovino. Puesto que han sido muchos los estudios realizados, en algunas ocasiones los resultados no resultan coincidentes. En este sentido, con el objetivo de confirmar si el efecto de las variantes de la kappa-caseína en diferentes rasgos productivos y composicionales de la leche es significativo, en 2020 se realizó un meta-análisis (Mahmoudi et al., 2020). En dicha publicación, se integraron los datos reportados por diferentes estudios literarios publicados hasta el momento, logrando agrupar un conjunto de 30.471 bovinos genotipados. De las 14 variantes identificadas, se decidió estudiar específicamente los alelos A y B, puesto que se observó que estos presentan las mayores frecuencias alélicas y se encuentran prácticamente en todas las razas de ganado bovino. Finalmente, el estudio concluyó que existe un efecto estadísticamente significativo de los genotipos del *CSN3*, sobre los rasgos de producción diaria de leche, rendimiento de la lactación, porcentaje de grasa y porcentaje de proteína en el ganado vacuno. Además, entre las dos variantes estudiadas, se pudo confirmar que el alelo B afecta positivamente al rendimiento lácteo diario, al porcentaje de grasa y al porcentaje de proteína en el ganado.

De esta forma, este meta-análisis confirma los resultados hasta ahora publicados sobre las variantes alélicas en la raza Frisona. Como se mencionaba con anterioridad, en esta raza las variantes más estudiadas son los alelos A, B y E, de las que los estudios coinciden en que la leche producida por vacas homocigóticas en el gen *CSN3* de genotipo BB, presenta mejores propiedades tecnológicas como menor tiempo de coagulación y mayor rendimiento quesero en comparación con las variantes A y E. Además, también señalan que los genotipos AA y el alelo A, son los polimorfismos que presentan mayor frecuencia en la raza Frisona. Finalmente, también afirman que incrementar la frecuencia del alelo B y genotipo BB mediante la selección basada en sus genotipos y alelos, supone la mejora de la leche en rasgos productivos y económicos. De esta forma, mediante el cruzamiento de animales con genotipos BB y AB, se logrará el incremento de animales con genotipos BB (Adamov et al., 2020; Volkandari et al., 2017).

No obstante, por otra parte, también es preciso considerar las interacciones entre los genes de los diferentes grupos caseínicos y con las proteínas séricas, concretamente con la β -caseína (o β -CN) y la β -LG. Se ha comprobado que para sus respectivos genes (*CSN2* y *LGB*), y en ambos casos, la variante B implica mejores aptitudes tecnológicas. Según estudios realizados en la raza Frisona, también existe una relación de los genes *CSN2* con la coagulación de la leche y los rendimientos de leche y proteína (Caroli et al., 2009). Por este motivo, en la selección genética del ganado, también es interesante tener en cuenta los haplotipos o genotipos compuestos por ambos genes. En concreto, se ha definido que aquellos animales que presentasen al menos un alelo B en ambos loci, el tiempo de coagulación de la leche y firmeza de la cuajada es mejor (Caroli et al., 2009). Esto tiene especial interés, debido a que la frecuencia de algunos alelos y haplotipos caseínicos como es el alelo B de la β -caseína, han experimentado una drástica reducción en algunos programas de cría de diferentes países. Esto se debe a que a partir de 1990, se comenzó a asociar la variante A¹ del gen *CSN2*, con ciertos efectos negativos (Chessa et al., 2020). Algunos estudios han reportado que la digestión de la proteína A¹ de la β -caseína puede conducir a la formación de un péptido bioactivo con actividades opiáceas, la β -casomorfina 7 (BCM-7), sugerida como posible causa de la intolerancia a la leche y de varias enfermedades en humanos (Sebastiani et al., 2020). Este hecho podría condicionar la selección directa del alelo A², asociado con una mayor digestibilidad de la leche, y suponer una pérdida de haplotipos que están asociados a buenas propiedades tecnológicas. Por este motivo, la selección de otros alelos en otros loci de la caseína, como la variante B de la kappa-caseína, podría compensar la pérdida de variabilidad genética y de los haplotipos asociados a buenas propiedades queseras (Chessa et al., 2020).

En definitiva, se puede concluir que el aumento de las frecuencias alélicas de las mutaciones con efectos favorables en la elaboración de quesos alberga interés, y que dicha selección es especialmente aplicable a las variantes de las kappa-caseína, dada su gran influencia en el cuajado de la leche. Por consiguiente, el alelo B para el gen *CSN3* y especialmente el genotipo homocigótico BB, puede ser empleado como marcador genético en programas de selección de ganado lechero, principalmente en sistemas basados en la transformación a derivados lácteos, ya que se ha comprobado que es el criterio de selección más favorable en términos económicos de producción. Además, dicha selección favorece a su vez a la disminución de la frecuencia de los alelos A y E, que contrariamente empobrecen las propiedades tecnológicas de la leche.

1.5. Técnicas de genotipado del ganado bovino empleadas para la mejora de las propiedades reológicas de la leche

Inicialmente el estudio de los polimorfismos genéticos de las proteínas lácteas, se basaba en la caracterización de sus estructuras primarias. Para ello, se seleccionaban muestras de leche de hembras en fase de lactación y se empleaban técnicas como electroforesis (separación de las proteínas según las diferencias de carga neta), isoelectroenfoques (electroforesis espacial en el que las proteínas migran a través de un gradiente de pH hasta alcanzar su punto isoeléctrico) o perfiles cromatográficos (especialmente, la Cromatografía Líquida de Alta Resolución o HPLC). No obstante, posteriormente, gracias al desarrollo de técnicas moleculares basadas en el ADN, fue posible conocer el genotipo de las lactoproteínas (Barroso et al., 2001).

Las primeras técnicas moleculares, se desarrollaron gracias al descubrimiento de las endonucleasas, unas enzimas de restricción producidas por bacterias que cortan los ácidos nucleicos en secuencias muy específicas de 4 a 8 pb. De esta forma, polimorfismos como el cambio de una base o una delección o inserción entre dos sitios de corte de las enzimas, alteran las dianas de restricción de la enzima empleada. Esto permite por tanto, detectar polimorfismos genéticos según el número y tamaño de fragmentos de ADN que se originan al digerirlo. Esta técnica se denomina Southern Blot y consiste en: la digestión del ADN, separación electroforética en agarosa de los fragmentos resultantes y transferencia de los fragmentos separados a membranas de nitrocelulosa o nylon (Southern Blotting), e hibridación con sondas moleculares radiactivas o enzimáticas que son complementarias al gen estudiado y que posteriormente se pueden visualizar por métodos autorradiográficos o colorimétricos (Sellick, 2002). La gran ventaja de esta técnica es que no es necesario conocer la secuencia del gen. No obstante, presenta los inconvenientes de que requiere grandes cantidades de ADN y de que su aplicación es preferible para detectar alteraciones a gran escala, no siendo adecuada para la estudio de mutaciones puntuales como los SNPs (Barroso et al., 2001; Sellick, 2002).

Posteriormente, gracias a la secuenciación del genoma, fue posible estudiar los diferentes polimorfismos de las variantes proteicas, su localización y naturaleza; lo cual permitió el desarrollo de nuevas técnicas moleculares. Estas permiten seleccionar los polimorfismos de interés para someterlos posteriormente a su secuenciación, partiendo de productos amplificados por PCR (Polymerase Chain Reaction), técnica posteriormente descrita. De esta forma, previamente se amplifica la región concreta de ADN, en la que se encuentran los polimorfismos objeto de estudio. Esto permite delimitar la zona de interés y partir de cantidades mínimas de ADN, que incluso puede estar muy degradado. Posteriormente, dependiendo de la técnica acoplada, se realiza la detección y/o identificación de los polimorfismos (Barroso et al., 2001; Koopae & Koshkoiyeh, 2014; Sellick, 2002):

- **SSCP** (Single-Strand Conformation Polymorphisms o Polimorfismo de Conformación de las Cadenas Simples). Para comprender dicha técnica es preciso señalar que el ADN,

cuando se encuentra en forma de cadena simple, tiende a doblarse y formar estructuras secundarias complejas. Estas son estabilizadas por medio de enlaces intramoleculares débiles, como los puentes de hidrógeno, y se denominan dímeros. Dado que estas estructuras son diferentes según la secuencia nucleotídica, dos moléculas que se distingan por una diferencia puntual (SNPs), forman diferentes dímeros. De esta forma, el fundamento de la técnica radica en la detección de diferencias presentes entre los fragmentos de ADN amplificados y de cadena simple, por medio de la detección de los dímeros formados. Esto es posible mediante su separación en poliacrilamida no desnaturizante, ya que migran de distinta manera, y su posterior detección mediante tinción con plata o el empleo de primers marcados. Finalmente, el uso de controles permite identificar los diferentes polimorfismos.

- **PCR-ARMS** (Amplification Refractory Mutation System o Sistema de mutación refractario a la amplificación por PCR). Esta es una variante de la PCR convencional, en la que se favorece a la amplificación preferente del alelo mutante del gen a amplificar. Para ello, en el diseño de los primers, cerca o en el propio extremo 3' de uno de los primers, se incluye la secuencia complementaria y específica del alelo mutante.
- **PCR-S/ASO** (Sequence/Alele Specific Oligonucleotides u Oligonucleótidos Específicos de la Secuencia/Alelo). Esta técnica consiste en hibridar oligonucleótidos específicos o sondas de cada alelo que se pretende detectar, a las secuencias previamente amplificadas por PCR. De esta forma, una diferencia puntual entre el oligonucleótido y la cadena molde, evita tal unión permitiendo identificar diferentes polimorfismos.
- **PCR-RFLPs** (PCR-Restriction Fragment Length Polymorphisms o Polimorfismo de Longitud de Fragmentos de Restricción). Esta técnica, se fundamenta en las mismas bases metodológicas que la técnica Southern Blot previamente descrita. No obstante, en este caso, se emplean las enzimas de restricción para digerir el producto de la PCR. Posteriormente, los fragmentos resultantes se separan electroforéticamente en geles de agarosa o de poliacrilamida, y se visualizan e identifican las diferentes bandas por medio de controles.
- **ACRS** (Amplification Created Restriction Sites o Sitios de restricción creados por amplificación). En algunos casos no es posible aplicar la técnica PCR-RFLPs, puesto que las mutaciones a detectar “rompen” la secuencia diana de las enzimas comerciales disponibles. Por ello, en estos casos se emplea esta variación de la técnica, en la que para la previa amplificación se diseñan primers en los que se introduce la diana de las enzimas de restricción para una o varias de las variantes a genotipar.

De entre todas las técnicas descritas, se comprueba que para el genotipado del ganado bovino en el locus del gen *CSN3*, se emplean las técnicas PCR-RFLP (Adamov et al., 2020; Cortés-López et al., 2012; Duifhuis-Rivera et al., 2014; Miluchová et al., 2018; Molee et al., 2011; Tyulkin et al., 2018; Volkandari et al., 2017), SSCP (Zambrano Burbano et al., 2010) y ARMS (Fonseca et al., 2013). Según los diferentes estudios publicados sobre las variantes proteicas de la kappa-caseína de la raza Frisona, se constata que la técnica más empleada es la PCR-RFLP. Esta ha sido universalmente utilizada para el genotipado rutinario de mutaciones conocidas, puesto que sus resultados son fácilmente visibles y la técnica no requiere una previa optimización de sus condiciones. Por otra parte, en comparación a la ARMS, no es preciso el diseño de primers para cada uno de los alelos a identificar. Y en cuanto a la técnica SSCP, la PCR-RFLP también ofrece mejores resultados ya que la sensibilidad de la SSCP es limitada (solo permite identificar entre el 60-80% de polimorfismos) (Barroso et al., 2001; Sellick, 2002). De hecho, existen estudios publicados que comparan ambas técnicas y advierten que la técnica SSCP no logra identificar

todas las variantes, mientras que la PCR-RFLP, permite una mejor detección de las mutaciones puntuales y por ello una mejor estimación de las medidas de diversidad genética a partir de frecuencias poblacionales (Carlos Solarte et al., 2012).

No obstante, la técnica PCR-RFLP también presenta ciertas limitaciones: sólo detecta aquellos cambios que crean o destruyen la diana concreta de las enzimas de restricción utilizada, las enzimas comerciales son muy costosas y la técnica consume mucho tiempo (Koopae & Koshkoiyeh, 2014; Sellick, 2002). Además, como se ha comentado, en ciertas ocasiones no puede emplearse y se debe aplicar la técnica ACRS cuyos costes son aún más elevados, ya que se requieren primers muy largos y geles de elevada resolución que permitan discriminar los fragmentos digeridos cuya diferencia de tamaño es muy reducida (20-25 pb) (Barroso et al., 2001). Debido a esto, en los últimos años es posible observar estudios que tratan de implementar técnicas con una sensibilidad de detección mayor y más eficientes, capaces de genotipar un elevado número de muestras de forma rápida y económica. Estas son por ejemplo la qPCR o quantitative real-time PCR (Kovalchuk et al., 2019; Vafin & Gilmanov, 2021) o chips o microarrays comercializados por empresas como el Illumina Beadchip (Chessa et al., 2020). De esta forma, se han encontrado resultados satisfactorios en la implementación de la PCR cuantitativa en tiempo real, la cual logra suprimir los procedimientos de la técnica RFLP-PCR como la escisión por endonucleasas y la posterior separación electroforética de los fragmentos generados, lo cual requiere mucho tiempo y encarece la técnica (Kovalchuk et al., 2019; Vafin & Gilmanov, 2021).

1.6. Técnica PCR cuantitativa en tiempo real o qPCR

La PCR cuantitativa en tiempo real es una variante de la PCR tradicional, desarrollada por Kary Mullis. La PCR es una técnica que permite realizar millones de copias de una secuencia específica de ADN, a través del proceso catalítico desempeñado por la enzima ADN polimerasa. Esta es una enzima termostable denominada “Taq ADN polimerasa” que permite la replicación exponencial del ADN molde mediante la repetición cíclica de tres etapas. Para ello, son requeridos además otros elementos: una pareja de primers o cebadores, desoxirribonucleótidos trifosfatados (dNTPs: adenina, timina, citosina y guanina), ión magnesio (Mg^{2+} , cofactor de la enzima), una solución amortiguadora o buffer y H_2O libre de nucleasas como disolvente de la reacción. Los primers, son secuencias de oligonucleótidos complementarios a los extremos de la secuencia que constituye el fragmento que se desea amplificar. Presentan un tamaño entre 15-25 pares de bases y deben presentar una cantidad de G-C inferior al 55%, para impedir la formación de productos inespecíficos como dímeros de primers. La pareja está compuesta por dos secuencias diferentes de primers, “forward” y “reverse”, que permiten la extensión del ADN en dirección 5'→3'. En cuanto a los dNTP's, estos corresponden a las bases nitrogenadas empleadas por la Taq polimerasa para la síntesis de las nuevas hebras de ADN (Tamay De Dios L et al., 2013).

La reacción PCR se realiza en equipos denominados termocicladores, que permiten ajustar los programas de temperatura-tiempo requeridos para realizar las tres etapas en las que se basa la reacción: desnaturalización, hibridación y extensión. La desnaturalización se realiza a 90-95°C, y permite la separación del ADN en dos cadenas que sirven de ADN molde para la síntesis de las nuevas cadenas. Posteriormente, en la hibridación, los cebadores se alinean al extremo 3' e hibridan con los extremos complementarios del fragmento de interés (amplicón). Para ello, se requiere una temperatura denominada temperatura de hibridación o Melting Temperature (T_m), la cual depende de los cebadores empleados y suele comprender entre los 50-60 °C. Por último, en la extensión o elongación realizada a 72°C (temperatura a la que la enzima es funcional), la Taq-polimerasa actúa sobre el complejo ADN molde-primers y empieza a agregar los dNTP's

complementarios en dirección 5' → 3'. Estas etapas se repiten un determinado número de ciclos, hasta lograr la amplificación de un número detectable de amplicones. Finalmente, estos son visualizados realizando una electroforesis en geles de agarosa y exponiéndolo a luz UV. La electroforesis permite la separación de los ácidos nucleicos de acuerdo a su tamaño y carga eléctrica. Para ello, se emplea un gel de agarosa dispuesto en una matriz sólida, en el que se genera un campo eléctrico y dado que los ácidos nucleicos presentan carga negativa proporcionada por el grupo fosfato, estos migran hacia el polo positivo produciéndose paralelamente la separación de los fragmentos. Por último, su visualización es posible gracias a otro compuesto presente en la matriz, que se une al ADN y emite una señal al ser expuesto a luz UV. Este permite visualizar los amplicones en forma de bandas, las cuales pueden identificarse con su fragmento correspondiente mediante el uso de un marcador molecular de un tamaño o número de pares de bases conocido (FAO, 2007, Part 4; Tamay De Dios L et al., 2013).

Por su parte, la PCR cuantitativa en tiempo real, se fundamenta en el mismo principio de la técnica tradicional descrita. No obstante, esta difiere en la forma de detección y análisis de los productos amplificados (Barroso et al., 2001), en lugar de realizarse la detección una vez finalizada la reacción, con la qPCR se detecta y mide la acumulación del producto amplificado en cada ciclo. De esta forma, empleando termocicladores equipados con módulos de detección de fluorescencia, es posible registrar la fluorescencia a medida que se produce la amplificación y, en consecuencia, detectar la cantidad de producto amplificado. Por ello, la principal ventaja que ofrece esta técnica frente a la convencional, es que permite estimar con precisión y alta sensibilidad el número de copias de ADN presentes en la muestra analizada (resultados cuantitativos), a diferencia que la PCR tradicional que solo permite determinar la presencia o ausencia de una secuencia (resultados cualitativos). Además, esta técnica conlleva un menor tiempo de experimentación y reduce la posible contaminación de las muestras, dado que no es necesario realizar la electroforesis en gel (BIO-RAD Laboratories, 2006).

La detección “a tiempo real”, es posible gracias a la inclusión de una molécula fluorescente en la reacción. Esta permite correlacionar el aumento de la señal fluorescente que emite, con la cantidad proporcional de ADN que se va acumulando en cada ciclo. Dichas sustancias químicas consisten en tintes de unión al ADN o sondas específicas marcadas con fluorescencia, clasificadas respectivamente como sistemas no específicos y específicos (BIO-RAD Laboratories, 2006; Tamay De Dios L et al., 2013).

Los fluoróforos no específicos son los denominados agentes intercalantes. Estos son moléculas que tienen afinidad por el ADN de doble cadena, que al unirse, se oxidan y generan una señal fluorescente. De esta forma, durante la PCR los cebadores amplifican la secuencia diana, y múltiples fluoróforos se insertan entre las bases del producto de doble cadena, se oxidan y emiten fluorescencia. Algunos ejemplos son SYBR Green I, Cyto, EvaGreen y los colorantes LC (LightCycler Dyes). Estos son sistemas no específicos, ya que se unen de forma indiscriminada a ADN bicatenario y no a secuencias concretas. Por este motivo, las sondas inespecíficas resultan versátiles y económicas, sin embargo, esto también puede suponer la aparición de falsos positivos en los resultados. Si los agentes intercalantes se unen a productos de amplificación incorrectos o a dímeros de primers, la fluorescencia detectada se asociaría incorrectamente a la secuencia de interés. Por ello, en caso de emplear este tipo de fluoróforos, es preciso analizar la especificidad de la reacción mediante el análisis de las curvas melting o de disociación (Integrated DNA Technologies (IDT), 2012; Tamay De Dios L et al., 2013). Estas curvas se generan tras finalizar la reacción de amplificación, realizando un incremento de la temperatura por intervalos. De esta forma, a medida que el ADN se desnaturaliza, la señal fluorescente disminuye. Posteriormente, el software realiza gráficos como el representado en la Figura 4, en los que se representa la

derivada negativa del cambio de la fluorescencia en función de la temperatura. Esto permite observar un pico característico en la temperatura melting del amplicón (T_m), y puesto que dependiendo de la longitud y composición nucleotídica del producto resultante los picos son diferentes, esto permite diferenciarlo de otros productos inespecíficos que desnaturalizan a diferentes temperaturas (BIO-RAD Laboratories, 2006).

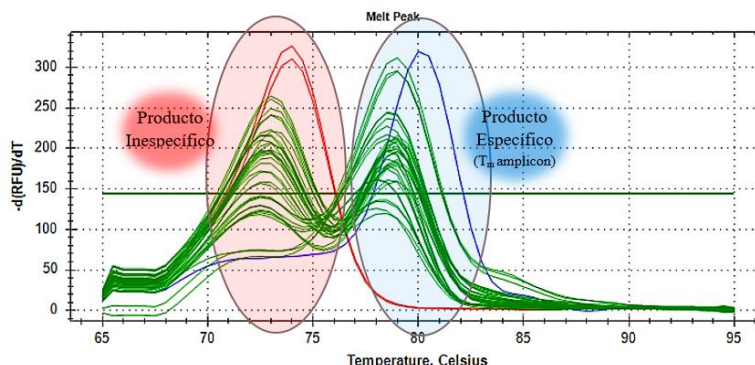


Figura 4. Curva de disociación de una qPCR realizada en laboratorio.

Por su parte, los sistemas específicos se basan en el principio denominado FRET (“Fluorescence Resonance Energy Transfer”), el cual consiste en la transferencia de energía fluorescente desde un donador fluorescente o “reporter”, a un aceptor o “quencher”. Concretamente, existen dos tipos de sondas, unas basadas en el método de hidrólisis y otras que funcionan por el método de hibridación.

Las sondas de hidrólisis son secuencias de nucleótidos complementarias a una secuencia específica del gen de interés, y que presentan un “reporter” en el extremo 5’ y un “quencher” en el extremo 3’. Su proximidad supone que se produzca el fenómeno FRET, es decir, la fluorescencia emitida por el “reporter” es absorbida por el “quencher” por lo que no puede ser detectada por el equipo. No obstante, en el momento en el que la sonda hibrida con su secuencia complementaria del ADN molde, se producen cambios conformacionales en sus estructuras, permitiendo que la actividad exonucleasa 5’-3’ de la Taq polimerasa rompa su unión. De esta forma, la hidrólisis separa ambos compuestos y libera la fluorescencia del donador, pudiendo ser capturada por el equipo. En cuanto a las sondas de hibridación, estas también presentan un “reporter” en el extremo 5’ y a un “quencher” en el extremo 3’. No obstante, en este caso no presentan una estructura lineal, sino que sus extremos son complementarios confiriéndole una forma de horquilla que permite la proximidad física del donador y aceptor. Por consiguiente, la fluorescencia no se detecta hasta que la sonda hibrida con la secuencia diana, ya que solo con la hibridación se produce la separación del aceptor y donador.

Las sondas comerciales de hidrólisis e hibridación más empleadas son respectivamente las sondas TaqMan y las molecular Beacons. En este caso, ambos tipos de sondas son más costosas que las no específicas, pero son más eficientes, ya que garantizan la especificidad de la reacción (IDT, 2012). Por este motivo, este tipo de sondas suele emplearse para realizar la discriminación de variantes alélicas. De esta forma, empleando diferentes colorantes fluorescentes como “reporters”, la hibridación preferencial y específica de una sonda al amplicón frente a otra sonda no emparejada y que por ello permanece apagada, permite detectar diferencias genotípicas tal y como se ejemplifica en la Figura 5 (BIO-RAD Laboratories, 2006).

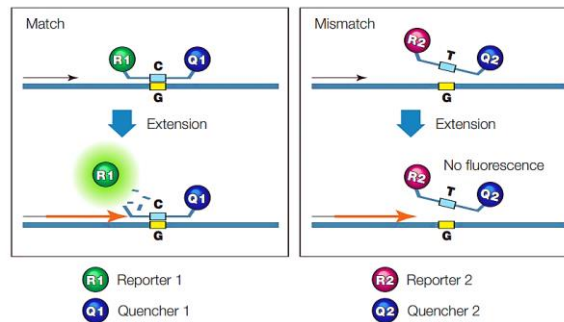


Figura 5. Mecanismo de discriminación alélica mediante el empleo de sondas TaqMan (BIO-RAD Laboratories, 2006).

En cuanto a la cuantificación del ADN, esto es posible por medio de los gráficos generados por el software de un equipo incorporado al termociclador (Tamay De Dios L et al., 2013). Uno de estos gráficos es la curva de amplificación, que refleja la fluorescencia acumulada tras cada reacción de amplificación, frente al número de ciclos de PCR. Durante el registro de la fluorescencia, tal y como se refleja en la Figura 6, es posible identificar hasta cuatro fases: Basal o “Baseline”, Exponencial o “Exponential”, Lineal o “Linear” y Meseta o “Plateau” (BIO-RAD Laboratories, 2006). No obstante, el gráfico también puede dividirse en una fase exponencial y en una no exponencial (IDT, 2012). Durante la fase exponencial la cantidad de ADN aumenta en cada ciclo, pero a medida que la reacción avanza, los componentes de la reacción se consumen y la reacción se ralentiza, dando lugar a la fase no exponencial. Durante el inicio del crecimiento exponencial, la fluorescencia se mantiene en niveles no detectables, pero llega un punto en el que el producto acumulado es suficiente y emite una señal detectable por el equipo. El número de ciclo en el que la fluorescencia supera el límite de detección o “Threshold line”, se denomina ciclo umbral (Ct) o ciclo de cuantificación (Cq), y es inversamente proporcional a la cantidad de ADN presente en la muestra. Valores inferiores de Cq indican que se ha requerido de un menor número de ciclos para amplificar el ADN hasta su detección, ya que este se encuentra en una mayor concentración (BIO-RAD Laboratories, 2006).

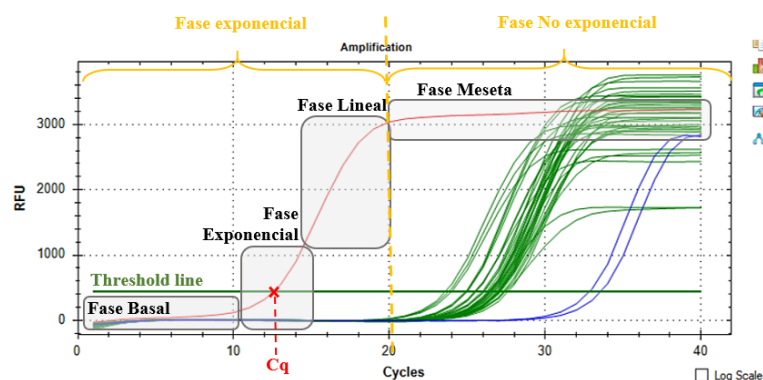


Figura 6. Gráfico de amplificación de una qPCR realizada en laboratorio.

Finalmente, para conferir sentido biológico al valor Cq obtenido, atendiendo a la finalidad del estudio realizado, es posible aplicar el método de cuantificación absoluta o relativa. El método absoluto, permite determinar la cantidad de ácido nucleico (número de copias o μg) por una cantidad determinada de muestra. Esto se realiza comparando los valores de Cq de las muestras, con una curva estándar elaborada mediante diluciones de un gen de referencia de cantidad conocida (BIO-RAD Laboratories, 2006). Por su parte, la cuantificación relativa, suele ser empleada para evaluar cambios en la expresión de genes reflejando la cantidad relativa del ácido nucleico frente a cantidades equivalentes de la muestra analizada y la de un gen de referencia (Tamay De Dios L et al., 2013).

2. OBJETIVOS

El **objetivo general** del presente trabajo es la puesta a punto de una técnica que permita el genotipado del ganado animal de forma sencilla y rápida. De esta forma, se pretende facilitar la selección de los genotipos o alelos del gen de la kappa-caseína, más favorables en términos productivos para la elaboración del queso. La técnica propuesta es la PCR en tiempo real mediante sondas específicas para la detección de los alelos A, B y E del gen de la kappa-caseína en vacuno de raza Frisona. Como se ha mencionado anteriormente, se puede asumir como hipótesis de trabajo que en la Frisona solo están presentes estos alelos, al menos de forma relevante. Para la consecución de la técnica, se proponen los siguientes **objetivos específicos**:

- Optimizar el método de extracción de ADN, determinando la matriz biológica de partida más favorable para la extracción de ADN en cantidad y de calidad, a través de folículos de pelo o de las células somáticas de la leche. Además, en el caso de la leche, evaluar la influencia de la temperatura de conservación de las muestras previa a la extracción.
- Diseñar y optimizar las condiciones de amplificación de la qPCR, estimando la concentración de cebadores y las condiciones que influyen en el funcionamiento de las sondas, más eficientes para la amplificación del gen y la discriminación de sus variantes.
- Validar el método de identificación de los alelos A, B y E mediante qPCR, en base a la eficiencia, sensibilidad, rango dinámico y repetibilidad del ensayo.
- Comprobar la aplicabilidad de la técnica desarrollada mediante el genotipado de ganado vacuno de la raza Frisona.

3. MATERIALES Y MÉTODOS

3.1. Obtención de las muestras

En el desarrollo de las diferentes experimentaciones que fundamentan el presente trabajo, se ha precisado de diferentes tipos de muestras, cuya utilización y fuente de obtención, se especifica a continuación. Previamente es preciso denotar que en todos los casos las muestras procedieron de la raza de ganado vacuno Holstein Frisona (especie *Bos taurus*).

Para optimizar el método de extracción empleado, evaluar el efecto de la temperatura en la conservación previa a la extracción y optimizar las condiciones de la qPCR, se emplearon muestras de leche cruda desnatada. Dichas muestras procedieron de establecimientos comerciales, y una vez recepcionadas, fueron etiquetadas y divididas en tres grupos muestrales, cada uno sometido a una temperatura de conservación diferente hasta la realización de la técnica de extracción: refrigeración a 4°C, congelación a -20°C y congelación a -80°C.

Las muestras empleadas para realizar la puesta a punto de la técnica qPCR mediante la validación final de sus resultados, fueron recogidas de la granja S.A.T. Etxeberri perteneciente al grupo S.A.T. Lacturale. En este caso, se realizó la recogida de muestras de pelo y leche, de animales previamente genotipados. Estas vacas fueron genotipadas mediante el chip de ADN EuroG MD dentro del proyecto ABIESNEA (210190016) del Programa de Desarrollo Rural de Navarra 2014-2021. La recogida de las muestras de leche se realizó por ordeño manual de cada uno de los individuos de la población muestral seleccionada, y se extrajo y seleccionó pelo con presencia de folículo piloso. Las muestras de leche y pelo fueron debidamente etiquetadas, transportadas y almacenadas, bajo condiciones de congelación y temperatura ambiente, respectivamente.

Por último, para optimizar las condiciones de la qPCR y validar los resultados, se precisaron muestras de referencia de las variantes A, B y E de la kappa-caseína. Dichas muestras de ADN sintético, empleadas como controles, fueron obtenidas de la casa comercial Integrated DNA Technologies (IDT).

3.2. Extracción de ADN a partir de células somáticas de leche y pelo

En la actualidad, existen numerosos métodos de extracción disponibles que permiten aislar ácidos nucleicos procedentes de diferentes matrices biológicas. No obstante, la calidad y cantidad de ADN extraído difiere de los fluidos o tejidos corporales empleados, lo cual es determinante para la posterior experimentación. Usualmente, la sangre ha constituido una de las fuentes de ADN más empleadas, sin embargo, su obtención implica ciertos riesgos como la exposición del animal a cierto estrés, que incluso puede conllevar un menor rendimiento productivo. Por este motivo, la leche ha sido valorada como un buen material de extracción, dado que su recogida es menos invasiva, es más sencilla, y requiere menos costes de servicio (Pokorska et al., 2016).

Concretamente, el ADN presente en la leche procede de las células somáticas del animal, las cuales abarcan un grupo diverso de células: leucocitos polimorfonucleares, macrófagos, linfocitos y un pequeño número de células epiteliales mamarias. Aunque el número de células en 1 ml de leche suele variar entre 2×10^4 y 2×10^5 células, su recuento puede verse afectado por diversos factores como la especie, la raza, la fase de la lactación, la higiene durante el ordeño, el estrés, el desarrollo de mastitis y las predisposiciones individuales (Pokorska et al., 2016). Por este motivo, pese a las ventajas que ofrece en la técnica de extracción, su bajo contenido y variabilidad instan a la búsqueda de otras matrices, tal y como es posible observar en otros muchos estudios en los que también se realizan muestreos a partir de tejidos como el pelo (Bartonova et al., 2012; Ionel et al., 2017; Rangel et al., 2017; Trakovickà et al., 2012). Por este mismo motivo, en el presente

trabajo se valora realizar una comparativa de la cantidad y calidad del ADN extraído procedente de las células somáticas de la leche y del pelo del animal. Por otra parte, es preciso recordar que el genotipado a partir de células somáticas, se reduce a hembras en fase de lactación. De esta forma, otra ventaja adicional que ofrece la extracción de ADN a partir de pelo, es la posibilidad de definir el genotipo para el gen *CSN3* en animales de cualquier sexo y edad.

3.2.1. Metodología empleada para la extracción

La extracción del ADN para ambos tipos de muestras, leche y raíces del pelo, se realizó mediante el kit comercial NucleoSpin Tissue (Macherey-Nagel, 2014). En este caso, se siguió el protocolo definido por el fabricante descrito a continuación y esquematizado en la Tabla 1. No obstante, previamente fue requerido el pretratamiento de las muestras, el cual difirió según la matriz biológica de partida:

- Pretratamiento de las muestras de leche, definido por Yap et al. (2020):









Para cada muestra de leche se emplearon dos muestras de 10 ml. Cada muestra se trasvasó a un tubo falcon de 15 ml y se sometió a una centrifugación de 4500 g (6682 rpm), a 4°C y durante 20 minutos. Posteriormente, con ayuda de una espátula esterilizada, se eliminó la nata que hubiera podido aparecer. A continuación, se eliminó el sobrenadante, y para resuspender los pellets, se añadieron 400 µl de tampón PBS a cada tubo. Seguidamente, ambos volúmenes se juntaron en un eppendorf de 2 ml, que posteriormente fue centrifugado con short spin (1 min) a 4500 g. De nuevo, se eliminó el sobrenadante, se adicionó 1 ml de PBS y se volvió a centrifugar con short spin a 4500 g (este paso se repitió una vez más). Una vez finalizado el pretratamiento, el pellet se mantuvo a -20°C hasta la realización de la extracción. Para ello, previamente se esperó a la descongelación de los pellets, y a continuación, se realizó la etapa de prelisis descrita en el protocolo del kit NucleoSpin Tissue (Tabla 1).

- Pretratamiento de las muestras de pelo, definido en el protocolo del kit NucleoSpin Tissue (Macherey-Nagel, 2014):

Para las muestras de pelo de cada individuo, se depositaron las raíces previamente cortadas en un eppendorf de 1,5 ml. A continuación, se realizó una prelisis de la muestra subdividida en dos etapas. En la primera se adicionaron 360 µl de Buffer T1, el doble de lo especificado en el kit, para asegurar que este cubriese todas las raíces de la muestra. Posteriormente, cuatro veces de forma sucesiva y en el mismo orden, se repitió la congelación de las muestras con nitrógeno líquido y la posterior descongelación en un termobloque a 56°C. Por último, en la segunda parte, se adicionaron 25 µl de Proteinasa K, se realizó la mezcla con vortex y se incubaron las muestras a 56°C durante toda la noche. Tras el pretratamiento, la extracción continuó con la fase de lisis descrita por el mismo protocolo (Tabla 1).

Una vez aplicado el correspondiente pretratamiento a las muestras, se realizó la extracción con el kit mencionado. Para realizar la extracción y aislamiento del ADN, este kit consta de un sistema de columnas con membranas de sílice y tubos colectores. De esta forma, empleando los diferentes reactivos del kit y ajustando las condiciones especificadas en el protocolo, se llevaron a cabo las diferentes etapas descritas en la Tabla 1. Finalmente, el ADN extraído y eluido en el buffer de elución, se conservó a -20°C.

Tabla 1. Protocolo de extracción del kit NucleoSpin Tissue (modificado de Macherey-Nagel, 2014).

| <i>Fases</i> | <i>Reactivos</i> | <i>Condiciones</i> | |
|--|---|---|--|
| 1. Pre-lisis | + 180 µl Buffer T1 + 25 µl de Proteinasa K Mezclar con vórtex | Incubación con agitación a 56°C durante 2 horas. |  |
| 2. Lisis | Mezclar con vórtex + 200 µL Buffer de lisis B3 Mezclar con vórtex | Incubación a 70°C durante 10 minutos. |  |
| 3. Ajuste de condiciones de unión del ADN | + 210 µL de etanol frío (96–100 %) Mezclar con vórtex | Se ensambla una Columna NucleoSpin Tissue y un tubo colector de 2 ml, y se transfiere la muestra a la columna. |  |
| 4. Unión del ADN | | La columna se centrifuga 1 minuto a 11000 g. Se desecha el eluido y se coloca un nuevo tubo colector. |  |
| 5. Lavado de las membranas | 1º Lavado + 500 µL Buffer BW | Centrifugación a 11000 g durante 1 minuto. Se desecha el eluido y se coloca un nuevo tubo colector. |  |
| | 2º Lavado + 600 µL Buffer B5 (B5 con etanol) | Se centrifuga 1 minuto a 11000 g. Se desecha el eluido y se coloca un nuevo tubo colector. Se repite este paso una segunda vez. |  |
| 6. Secado de las membranas | | Centrifugación de 1 minuto a 11000 g y se coloca la columna en un eppendorf de 1,5 ml. |  |
| 7. Elución del ADN | + 50 µL Buffer elución (BE precalentado a 70 °C) | Se incuba 1 minuto con la tapa abierta y se centrifuga 1 minuto a 11000 g. |  |

3.2.2. Diseño experimental planteado para la optimización de la extracción

En la Tabla 2, se esquematiza el diseño experimental planteado para definir la matriz biológica más conveniente, para la obtención de una cantidad óptima de ADN de calidad. Como se ha comentado, la extracción de ADN a partir de células somáticas implica una cantidad considerable de leche para lograr una cantidad y calidad aceptables, por lo que la extracción a partir de las raíces de pelo podría constituir una buena alternativa. Por otra parte, en la búsqueda de la optimización de la extracción a partir de leche, la temperatura de conservación de las muestras se ha planteado como una posible fuente de variabilidad en el rendimiento y pureza del ADN extraído. Para ello, se plantean condiciones habitualmente empleadas en el laboratorio para la conservación de muestras. De esta forma, finalmente se definieron cuatro niveles para el factor de la extracción: leche a 4°C, leche a -20°C, leche a -80°C y pelo.

Por último, para evaluar el efecto de la temperatura y del tejido o fluido empleado en la extracción, se han propuesto cuatro variables en las que se puede parametrizar dicho efecto. Por un lado, la cantidad del ADN extraído puede estimarse mediante la cuantificación de la concentración, y la calidad, mediante la estimación de los ratios de absorbancia entre diferentes longitudes de onda (ambas metodologías se detallan en el apartado 3.3). Además, también es posible evaluar su influencia en la capacidad de amplificación del ADN extraído (*Amplificability*), por medio de qPCR. Las muestras sometidas a extracción pueden contener inhibidores de la PCR que pueden comprometer la capacidad de amplificación del ADN. Por ejemplo, en la leche, la presencia de altos niveles de grasas, proteínas y calcio puede mermar la eficacia de la amplificación (Pirondini et al., 2010). De esta forma, por medio del valor Cq obtenido en la qPCR, es posible evaluar al mismo tiempo la calidad y cantidad del ADN extraído. Puesto que el valor de Cq es inversamente proporcional a la cantidad de ADN presente en la muestra amplificada, si la cuantificación de la concentración (ng/µl) guarda correlación con los valores Cq obtenidos, se puede asumir la ausencia de inhibidores que comprometan la correcta amplificación del ADN extraído.

Tabla 2. Matriz del diseño experimental planteado para la optimización de la extracción.

| <i>Factor Extracción</i> | | <i>Variables dependientes</i> | | | |
|--------------------------|------------------------------------|-------------------------------|------------------------------|------------------------------|-------------------------|
| Matriz biológica | Condiciones de conservación | Concentración (ng/μl) | A _{260/280} (nm) | A _{260/230} (nm) | Amplificability (Cq) |
| Leche | Refrigeración a 4°C | | | | |
| | Congelación a -20°C | | | | |
| | Congelación a -80°C | | | | |
| Pelo | Tª ambiente | | | | |

Para la realización de este estudio se decidió emplear muestras de leche cruda desnatada obtenidas de establecimientos comerciales y pelo de las vacas genotipadas de la granja S.A.T. Etxeberri. Las muestras de leche de las vacas genotipadas no fueron empleadas en este ensayo dado que su cantidad disponible era limitada, su utilización sería más importante en la puesta a punto de la técnica qPCR y, además, no influye en la evaluación de la extracción.

Para lograr discernir si el efecto de la temperatura de conservación presenta significancia en la extracción de ADN a partir de la leche, se realizó el cálculo del tamaño muestral (n) en base a la ratio $A_{260/280}$. No se tuvo en cuenta la concentración de la muestra de ADN, ya que este no es un aspecto tan determinante para su posterior amplificación mediante qPCR. Para ello, se hizo uso de la herramienta “pwr.t.test” del paquete estadístico R, en el que se realizó una prueba t-student de comparación de medias a dos colas con los valores de significación del 95% y una potencia del 70%. Además, se definió el valor “d” (diferencia entre las medias dividida por la desviación estándar (SD)) como 1, siendo la diferencia a detectar de 0,2 y la SD de 0,2. Este valor se estableció en base a datos bibliográficos consultados, en los que el valor de SD de los métodos de extracción a partir de células somáticas de leche era de 0,1 (Pokorska et al., 2016) a 0,3 (Wassermann, 2020). Finalmente, en base a todos los parámetros indicados, se determinó un tamaño muestral de 14 muestras para cada uno de los tres niveles. Por el contrario, en el caso del pelo, solo se pudo disponer de 12 muestras procedentes de dos muestras de pelo de cada uno de los seis posibles genotipos para los tres alelos estudiados (AA, AB, BB, AE, BE, EE).

De esta forma, tras la conservación de las muestras durante dos días a su temperatura correspondiente, se realizó la extracción de las muestras de acuerdo a la técnica previamente descrita, y se cuantificó y evaluó la pureza del ADN extraído, de acuerdo a lo descrito en el apartado 3.3 de la metodología. Posteriormente, se realizó la amplificación del ADN extraído por medio de la PCR en tiempo real. Cabe denotar, que de forma previa a la realización del presente estudio sobre la kappa-caseína, en el laboratorio se realizó la puesta a punto de la metodología de extracción de leche y de las condiciones de la qPCR para el gen *CSN2* o gen de la β-caseína. Dado que estos resultados condicionan la posterior optimización y puesta a punto de la qPCR para el gen de la kappa-caseína, se decidió evaluar la extracción previa mediante la amplificación del gen de la β-caseína, y así posteriormente, aplicar los resultados más favorables en el desarrollo de la técnica.

Por consiguiente, de acuerdo a las cantidades y concentraciones fijadas en el previo estudio mencionado, se realizó la preparación de los diferentes reactivos para completar un volumen final de reacción de 10 μl: 5 μl de Sybr Green (2x), 0,3 μl de Primer Forward del gen *CSN2* (10 μM), 0,3 μl de Primer Reverse del gen *CSN2* (10 μM), 2 μl de muestra y el volumen restante de agua libre de DNAsas. La reacción de amplificación se realizó en placas de 96 pocillos (Multiplate™ 96-Well PCR Plates, BIO-RAD), en el termociclador CFX96 Touch Real-Time PCR Detection System (BIO-RAD), con el mismo programa seleccionado por el estudio: primera etapa de desnaturalización y activación de la polimerasa a 95°C durante 30 segundos, 40 ciclos de amplificación conformados por una primera desnaturalización a 95°C durante 5 segundos y una

etapa de annealing o extensión a 60°C durante 30 segundos, y una curva de melting en el que la temperatura aumenta 0,5°C cada 5 segundos hasta alcanzar de nuevo 95°C. Cada una de las muestras se amplificó por triplicado, como blanco se empleó agua libre de RNAsas, y como control positivo, una muestra extraída del alelo A² del gen *CSN2* que contenía 3,67x10⁶ copias/μl.

Finalmente, mediante los datos proporcionados por el software del termociclador, se obtuvieron los valores C_q de todas las muestras amplificadas por triplicado. Además, puesto que la qPCR se realizó con el fluorocromo inespecífico SYBR Green, para confirmar la ausencia de productos inespecíficos como los dímeros de cebadores, se realizó el análisis de las curvas de melting o curvas de disociación. De esta forma, mediante la visualización de un solo pico, se pudo asegurar la única amplificación del producto deseado, y consecuentemente, confirmar la especificidad de la reacción.

Por último, el análisis estadístico de los datos obtenidos se realizó mediante el paquete estadístico IBM SPSS. Para ello, primero se estimó el valor de las medias y desviaciones típicas de las cuatro variables fijadas (concentración de ADN, ratios de absorbancia a diferentes longitudes de onda (A_{260/280} y A_{260/230}) y *Amplificability*), para los cuatro grupos muestrales diseñados (leche a 4°C, leche a -20°C, leche a -80°C y pelo). A continuación, por medio del programa estadístico mencionado, se evaluó el efecto principal y general del factor “extracción” en las cuatro variables, mediante un análisis de la varianza o ANOVA simple. Posteriormente, se realizó una prueba multivariante o “prueba post-hoc” mediante el método Tukey HSD (Honestly significant differences), para evaluar el efecto individual y el efecto de la interacción entre los cuatro niveles de extracción para las cuatro variables planteadas. En todos los casos, el nivel de significación estadística se fijó en 0,05 (nivel de significación del 95%).

3.3. Cuantificación y pureza del ADN extraído

Tras el proceso de extracción y aislamiento, es fundamental valorar la pureza y rendimiento del ADN extraído. Esto se debe a que es preferible que las muestras sean homogéneas, tanto en calidad como cantidad, para prevenir posibles errores en la posterior experimentación e interpretación de resultados. Por lo tanto, esto se realizó tanto en el estudio de la extracción, como en la puesta a punto de la técnica qPCR para la kappa-caseína. Para ello, se empleó el espectrofotómetro NanoDrop (Thermo Scientific).

La cuantificación del ADN (ng/μL) se realizó a través de la medición de la absorbancia a 260 nm, adicionando 1 μl de cada muestra, limpiando el equipo entre muestra y muestra, y calibrando el equipo con el buffer BE en el que se eluye el ADN al finalizar su extracción. En cuanto a la evaluación de la pureza, esta se estimó en base a la relación de las lecturas de absorbancia de Ultravioleta (UV) a las longitudes de onda de 230, 260 y 280 nm. De esta forma, el valor ideal para la ratio A_{260/280}, se encuentra en valores comprendidos entre 1,8 y 2; y para la ratio A_{260/230}, en torno a valores de 2,2. Estos valores se asocian al ADN, por lo que ratios más bajos podrían indicar la contaminación por presencia de compuestos orgánicos, como fenol o proteínas (Usman et al., 2014).

3.4. Amplificación del gen *CSN3* mediante PCR en tiempo real (qPCR)

Tal y como se ha definido en los objetivos, con el presente trabajo se pretende desarrollar una técnica que permita genotipar el ganado vacuno de forma eficiente, identificando los alelos A, B y E de la kappa-caseína. Para ello, primero se debe realizar el diseño de la pareja de cebadores que permita amplificar la región del gen en la que se encuentran dichos polimorfismos, así como diseñar las sondas específicas que permitan identificarlos. Por otra parte, también se debe realizar el diseño de controles positivos de las tres variantes, con los que poder verificar la ausencia de

falsos negativos, además de elaborar curvas de calibrado para poder realizar cuantificaciones absolutas. Por último, se debe ajustar las condiciones de la qPCR, hasta lograr optimizar el funcionamiento de los cebadores y sondas diseñados. De esta forma, seguidamente se describe cada uno de los pasos realizados para la puesta a punto de la técnica desarrollada.

3.4.1. Diseño de cebadores y sondas

Los cebadores y sondas fueron diseñados a través de la herramienta “PrimerQuest Tool”, disponible en el sitio web de la casa comercial IDT (IDT, 2021c), que posteriormente se encargó de sintetizarlos. Para su diseño, inicialmente se realizó la búsqueda bibliográfica de la secuencia del gen de la kappa-caseína o gen *CSN3* para la especie *Bos Taurus*. Concretamente, fue utilizada la base de datos de secuencias de ADN “GenBank” (NCBI, 2021), con la que también se identificaron las posiciones de las mutaciones en la secuencia de los tres alelos. Tal y como refleja la Figura 3 en el apartado 1.3 de la introducción, el alelo B presenta dos mutaciones con respecto al alelo A, mientras que el alelo E solo una. De esta forma, se pudo identificar un fragmento del gen en el que, con relativa proximidad, se encuentra uno de los polimorfismos del alelo B (posición 13104) y la mutación del alelo E (13124). Por consiguiente, a partir de la información obtenida, se pudo realizar el diseño de la pareja de primers que permiten amplificar dicha región del gen *CSN3* (secuencia del amplicón representada en la Tabla 3). Además, también se diseñaron cuatro sondas específicas TaqMan, que permiten discriminar las variantes alélicas de cada posición: en la posición 13104, el alelo A y E (Adenina) del alelo B (Citosina), y en la posición 13124, el alelo A y B (Adenina) del alelo E (Guanina). Para ello, se siguieron las especificaciones recomendadas por IDT para el diseño de cebadores y sondas. Estas usualmente recogen los valores ideales de longitud, temperatura melting (T_m) y composición de bases con respecto al contenido de Guanina y Citosina (Contenido GC); para los cebadores, las sondas y el amplicón.

La temperatura melting o temperatura de fusión del ADN (T_m), es la temperatura en la que el 50% de las hebras de ADN se encuentran separadas. Durante la etapa annealing, los cebadores y sondas hibridan con sus dianas formando dúplex cortos, cuya estabilidad depende de dicha temperatura. Por este motivo, la T_m es un parámetro a controlar importante, ya que condiciona la unión de los primers y sondas. Otro aspecto importante es el porcentaje del número de ácidos nucleicos de bases Guanina y Citosina (Contenido GC), ya que permite asegurar que los cebadores y sondas sean específicos del ADN diana. Además, también garantiza que no sean complementarios entre sí y/o con otras secuencias, lo cual podría conllevar la formación de estructuras secundarias. Por último, la longitud de los primers y cebadores se estima conforme a los anteriores parámetros, y en cuanto a la región a amplificar o amplicón, el único requisito es una longitud suficientemente larga que permita la hibridación de las sondas y cebadores (70-150 pb).

Cuando las sondas utilizadas para la discriminación alélica son sondas TaqMan, las directrices aplicables difieren de las recomendaciones generales sugeridas por Livak (1999) y recogidas en las guías IDT (BIO-RAD Laboratories, 2006; IDT, 2012): sondas con una T_m de 65 a 67°C y una longitud de 23-26 bases, y cebadores con una T_m de 58 a 60°C. Por consiguiente, en este caso tal y como sugiere la casa comercial, es preferible que las sondas para la discriminación alélica sean más cortas, lo cual supone valores inferiores de T_m . Además, es imprescindible que las sondas presenten una alta especificidad con su variante alélica coincidente, para así solo unirse a la diana de su alelo específico (BIO-RAD Laboratories, 2006). De esta forma, en base a todos los criterios mencionados, en la Tabla 3 se indica la región a amplificar, la pareja de primers y los dos pares de sondas diseñados. Además, en la tabla 4 se indican las características de cada uno de ellos.

Tabla 3. Secuencia del amplicón, primers y sondas específicas; diseñados para la puesta a punto de la qPCR para la identificación de las variantes alélicas A, B y E de la kappa-caseína.

| | | | |
|---------------------------------------|------------------------|---------|---|
| Amplicón* | | | 5'- AAGTACACCTACCACCGAAGCAGTAGAGAGCACTGTA GCTACTCTAGAAG[A/C]TTCTCCAGAAGTTATTGAG[A/G]G CCCACCTGAGATCAACACAGTCCAAGTTAC – 3' |
| Forward Primer | | | 5'- AAGTACACCTACCACCGAAG – 3' |
| Reverse Primer | | | 5'- GTAACCTGGACTGTGTGATCTC – 3' |
| Sonda TaqMan posición 13104 | Alelo A y E Adenina | 13104_A | 5'- AGAA+G+A+TT+CT+C+CA – 3' |
| | Alelo B Citosina | 13104_C | 5'- CTA+CT+CTA+GAA+G+C+TT – 3' |
| Sonda TaqMan posición 13124 | Alelo A y B Adenina | 13124_A | 5'- TT+A+TTGA+G+A+GC+CC – 3' |
| | Alelo E Guanina | 13124_G | 5'- TAT+T+GA+G+G+GC – 3' |

*En sombreado gris se destacan las secuencias del par de primers, y entre corchetes, las posibles mutaciones de cada posición.

Tabla 4. Características del amplicón, primers y sondas específicas, diseñados para la puesta a punto de la qPCR para la identificación de las variantes alélicas A, B y E de la kappa-caseína.

| | Amplicón | Forward Primer | Reverse Primer | Sonda TaqMan posición 13104 | | Sonda TaqMan posición 13124 | |
|----------------------|----------|-------------------|-------------------|---------------------------------------|---------|---------------------------------------|---------|
| | | | | 13104_A | 13104_C | 13124_A | 13124_G |
| <i>Longitud</i> | 102 pb | 20 pb | 23 pb | 13 pb | 15 pb | 13 pb | 10 pb |
| <i>T_m</i> | - | 53,8°C | 53,3°C | 53,6°C | 56,1°C | 56,2°C | 55,5°C |
| <i>Contenido GC</i> | - | 50% | 43,25% | 38,5% | 40% | 46,2% | 50,0% |

Por otra parte, para asegurar un mayor éxito en el diseño de la técnica, también se realizó un análisis *in silico* por medio de otras dos herramientas. Por un lado, la herramienta “BLAST” (NCBI, 2021a), que permite garantizar la especificidad de los primers y sondas diseñados. Con esta se aseguró la ausencia de complementariedad con otras secuencias, lo cual impide posibles amplificaciones fuera de la región definida. Por otro lado, se empleó la herramienta IDT “OligoAnalyzer”, también disponible en su página web (IDT, 2021b), con el objetivo de buscar sitios potenciales para la formación de estructuras secundarias. En este caso, se advirtió la posible formación de una horquilla o “Self-Dimer” en la sonda de la posición 13104 para el alelo B, y un heterodímero entre las sondas de las dos variantes alélicas de la posición 13124 (Tabla 5). No obstante, la propia casa comercial indicó que debido a que el valor de su estabilidad termodinámica definida por la energía libre de Gibbs se aproxima a -9 kcal/mol, estos no llegarían a formarse y no influirían negativamente en el rendimiento del ensayo. En cualquier caso, esto se tuvo en consideración durante el análisis de los gráficos de amplificación, para poder asegurar la validez de los resultados y el correcto funcionamiento de las sondas.

Tabla 5. Estructuras secundarias detectadas con la herramienta “OligoAnalyzer” en el diseño de la qPCR para la identificación de las variantes alélicas A, B y E de la kappa-caseína.

| Estructuras secundarias | | | Energía libre de Gibbs (ΔG°) |
|--|----|-------------------|--|
| <u>Self-Dimer o horquilla</u> 13104_C | 5' | CTACTCTAGAAGCTT | -10,23 kcal/mol |
| | 3' | TTTTCGAAGATCTCATC | |
| <u>Hetero-Dimer</u> 13124_A y 13124_G | 5' | TTATTGAGAGCCC | -9,28 kcal/mol |
| | 3' | CGGGAGTTAT | |

Por último, para completar el diseño de la técnica, se debe escoger un colorante fluorescente. Puesto que su elección depende de su compatibilidad con el equipo empleado, se consultó en la guía IDT los colorantes compatibles con el termociclador CFX96 Touch Real-Time PCR Detection System empleado en el laboratorio. Este presenta canales capaces de detectar los

espectros de emisión de los colorantes: Fam, Hex, Texas red, Cy5, Quasar 705. Por otra parte, también se tuvo en cuenta que para poder aplicar dicha técnica en el genotipado animal, lo más oportuno sería que en una misma reacción se pudiese identificar las dos variantes alélicas presentes (necesario para el genotipado de animales homocigóticos y heterocigóticos para las tres variantes estudiadas del gen *CSN3*). Esto implica que las cuatro sondas deben incluirse en una misma reacción PCR, denominada PCR en multiplex, lo cual también supone que cada sonda debe presentar un fluorocromo diferente para que el equipo sea capaz de diferenciarlos mediante su detección en diferentes canales. No obstante, debido a la complejidad del diseño de sondas marcadas con cuatro colorantes diferentes y siguiendo las recomendaciones de la casa comercial, se optó porque solo las sondas de la misma posición presentasen fluorocromos diferentes: los fluoróforos FAM y HEX, respectivamente para las sondas 13104_A y 13104_C; y Cy5 y FAM, respectivamente para las sondas 13124_A y 13124_G. Así, en lugar de una multiplex se realizaría una duplex, incluyendo en cada reacción las dos sondas de la misma posición. Por consiguiente, para poder genotipar las muestras analizadas, cada muestra debería someterse a dos amplificaciones, una para cada pareja de sondas.

3.4.2. Diseño de muestras de referencia comerciales de ADN sintético

Por otra parte, al igual que en cualquier otra técnica experimental, se precisa de controles positivos para poder verificar la credibilidad de los resultados. En este caso, de nuevo se solicitó a la casa comercial IDT la síntesis de tres controles, uno para cada una de las variantes del gen *CSN3* (Tabla 6). Esta casa ofrece fragmentos génicos denominados “gBlocks”, que son fragmentos de doble cadena de un número de copias conocido. De esta forma, estos controles fueron empleados como muestras de referencia para prevenir la aparición de falsos negativos en los resultados, ya que permiten comprobar si la amplificación ausente del ADN diana, se debe o no a la presencia de inhibidores. Por otra parte, estos controles exógenos son de gran utilidad ya que pueden ser empleados como estándares de concentración conocida, y por tanto, ser utilizados en la elaboración de curvas de calibración que permiten realizar cuantificaciones absolutas, siendo indispensables para la validación de la técnica desarrollada.

Tabla 6. Muestras de referencia sintéticas de las variantes A, B y E de la kappa-caseína empleadas como controles.

| <i>Variantes</i> | <i>Secuencia genética</i> | <i>Características</i> |
|------------------|---|--|
| Alelo A | 5'- ATTGCTAGTGGTGAGCCTACAAGTACACCTACCACCG AAGCAGTAGAGAGCACTGTAGCTACTCTAGAAGATTCTC CAGAAAGTTATTGAGAGCCACCTGAGATCAACACAGTCC AAGTTACTTCAACTGCAGTCTAAAACTCT - 3' | <u>Tamaño</u> 145 pb |
| Alelo B | 5'- ATTGCTAGTGGTGAGCCTACAAGTACACCTACCATCG AAGCAGTAGAGAGCACTGTAGCTACTCTAGAAGCTTCTC CAGAAAGTTATTGAGAGCCACCTGAGATCAACACAGTCC AAGTTACTTCAACTGCAGTCTAAAACTCT - 3' | <u>Concentración</u> 10ng/μl |
| Alelo E | 5'- ATTGCTAGTGGTGAGCCTACAAGTACACCTACCACCG AAGCAGTAGAGAGCACTGTAGCTACTCTAGAAGATTCTC CAGAAAGTTATTGAGGGCCACCTGAGATCAACACAGTCC AAGTTACTTCAACTGCAGTCTAAAACTCT - 3' | <u>Peso Molecular</u> 11,18 fmol/ng <u>Nº copias</u> 6,7x10 ¹⁰ copias/μl |

3.4.3. Optimización de las condiciones de amplificación mediante qPCR

3.4.3.1. Optimización de la concentración de cebadores

Para realizar la selección de las condiciones más óptimas para la amplificación e identificación de las tres variantes alélicas, primero se debe identificar la concentración de cebadores más favorable. Para ello, se realizó el diseño de una matriz conformada por las combinaciones resultantes de tres concentraciones de los cebadores, forward y reverse, y tres concentraciones de

ADN (Tabla 7). Para realizar la selección de las concentraciones de cebadores se tomó como referencia anteriores trabajos publicados, en los que el objeto de estudio también era la kappa-caseína (Kovalchuk et al., 2019; Vafin & Gilmanov, 2021). De esta forma, con el diseño experimental planteado, se pretende estimar la concentración de los cebadores diseñados que permite una amplificación óptima del gen *CSN3* independientemente de la cantidad de ADN presente en la muestra.

Tabla 7. Matriz de cebadores para la optimización de la amplificación del gen *CSN3*.

| Concentración para cada pareja de cebadores (FW y RV) | Cantidad de ADN | | | | | | | | |
|---|-----------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| | 30 ng | | | 45 ng | | | 60 ng | | |
| | 0,2 μ M | 0,3 μ M | 0,4 μ M | 0,2 μ M | 0,3 μ M | 0,4 μ M | 0,2 μ M | 0,3 μ M | 0,4 μ M |

Para no limitar en posteriores ensayos la cantidad disponible de ADN extraído de leche o pelo de las vacas genotipadas, en este ensayo también se decidió emplear la leche cruda desnatada comercial. Concretamente, se emplearon las muestras extraídas de la leche previamente conservada a -80°C , dado que como se detalla posteriormente en el apartado 4.1 de resultados, la extracción a partir de leche conservada a menores temperatura resultó ser más óptima. De esta forma, las reacciones se prepararon conforme a los volúmenes dispuestos en la Tabla 8. No obstante, previamente para poder disponer de la concentración deseada de ADN mediante la adición de los volúmenes indicados, se preparó un pool de las 14 muestras analizadas. Para este pool se estimó una concentración media de 21 ng/ μ l, en base a la media total de las cantidades cuantificadas en cada muestra con el espectrofotómetro. Posteriormente, el pool se diluyó a 15 ng/ μ l para que al añadir 2 μ l, 3 μ l y 4 μ l del pool diluido, la cantidad de ADN adicionado fuese 30 ng, 45 ng y 60 ng, respectivamente.

Tabla 8. Volumen de los reactivos (μ l) para la optimización de la concentración de los cebadores para la amplificación del gen *CSN3* mediante PCR en tiempo real.

| Reactivos | 30 ng | | | 45 ng | | | 60 ng | | |
|-----------------------------|-------|-----|-----|-------|-----|-----|-------|-----|-----|
| SYBR Green (2x) | 5 | | | 5 | | | 5 | | |
| Primer forward (10 μ M) | 0,2 | 0,3 | 0,4 | 0,2 | 0,3 | 0,4 | 0,2 | 0,3 | 0,4 |
| Primer reverse (10 μ M) | 0,2 | 0,3 | 0,4 | 0,2 | 0,3 | 0,4 | 0,2 | 0,3 | 0,4 |
| Agua | 2,6 | 2,4 | 2,2 | 1,6 | 1,4 | 1,2 | 0,6 | 0,4 | 0,2 |
| ADN | 2 | | | 3 | | | 4 | | |
| Volumen total (μ l) | 10 | | | 10 | | | 10 | | |

Al igual que en el resto de ensayos, se empleó el termociclador CFX96 Touch Real-Time PCR Detection System, para el que se seleccionó el programa descrito en la Figura 7. Además, también se emplearon las placas de 96 pocillos, en las que, para cada una de las nueve combinaciones, se añadieron las muestras por triplicado y los correspondientes blancos.

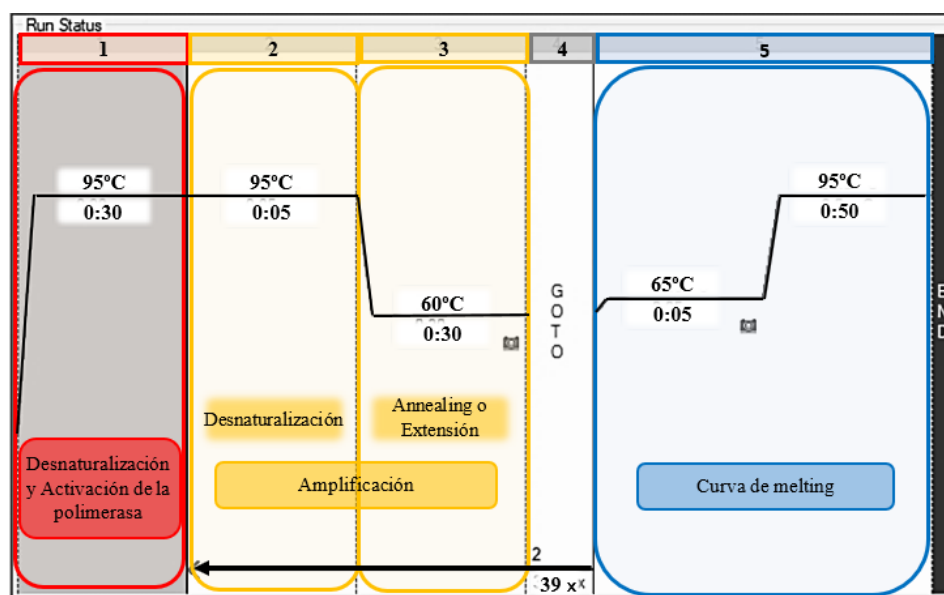


Figura 7. Programa de amplificación (Temperatura-Tiempo) para la amplificación del gen *CSN3* con el agente SYBR Green.

Como se puede comprobar en la Figura 7, puesto que se empleó la sonda inespecífica SYBR Green, en el programa se incluyó una curva de melting para posteriormente poder asegurar la ausencia de productos inespecíficos o dímeros de primers. De esta forma, tras analizar dichos gráficos y asegurar la validez de los resultados, se procedió a realizar su análisis estadístico mediante el paquete IBM SPSS. En este caso, para evaluar la concentración más óptima de cebadores, se estudió su efecto en dos variables, el valor C_q y la magnitud absoluta de las unidades relativas de fluorescencia o “relative fluorescence units” (RFU). Por un lado, menores valores de C_q sugieren una mejor amplificación del gen de interés, por lo que aquellas concentraciones que logran amplificar el gen en un menor número de ciclos, son más eficientes. Por otra parte, mayores niveles de fluorescencia alcanzados al final de la amplificación (“End RFU”), también sugieren una mejor amplificación del producto. De esta forma, para ambas variables inicialmente se realizó una ANOVA simple, y posteriormente, para evaluar el efecto individual y el contraste entre las diferentes combinaciones de concentraciones cebadores-muestra, se realizó la prueba multivariante mediante el método Tukey HSD. De nuevo se fijó un nivel de significación del 95%.

3.4.3.2. Optimización del funcionamiento de las sondas

Una vez estimada la concentración necesaria de cebadores en la reacción, se debe realizar el ajuste de los diferentes parámetros experimentales que condicionan el funcionamiento de las cuatro sondas diseñadas. De esta forma, se deben estimar las mejores condiciones para lograr la discriminación alélica de las tres variantes estudiadas del gen *CSN3* en reacciones duplex. Para ello, se realizó la consecución de una serie de pruebas en las que se modificaron alternativamente diferentes variables. De esta manera, mediante el método ensayo y error se fueron ajustando de forma empírica y progresiva, las diferentes condiciones hasta optimizar el funcionamiento de cada sonda. Esta metodología se aplicó principalmente, debido al precio de las sondas que limitó el número de ensayos posibles.

En la Tabla 9, se especifican las condiciones experimentales de las diferentes variables estudiadas en cada una de las pruebas realizadas. Tal y como se comprueba, antes de estudiar el funcionamiento de las sondas en duplex, se realizaron tres primeras pruebas en singleplex que permitieron estudiar su funcionamiento individual y fijar ciertos parámetros operacionales.

Tabla 9. Conjunto de pruebas realizadas para la optimización de las cuatro sondas: FAM_13104_A, HEX_13104_C, Cy5_13124_A y FAM_13124_G.

| Pruebas | Muestras de ADN | Sondas | | Programa termociclador |
|---------|--|---|---|--|
| Nº | Tipo de muestras de ADN | Concentración | Sondas empleadas | Etapas de annealing /extensión |
| 1 | Pool del ADN extraído de la leche conservada -80°C ¹ | 0,1 µM y 0,05 µM | En Singleplex las cuatro sondas | 60°C durante 30'' |
| 2 | Pool de las muestras de referencia de las tres variantes ² | 0,1 µM y 0,07 µM | En Singleplex las cuatro sondas | 60°C durante 30'' |
| 3 | Muestras de referencia adicionadas para su sonda específica ³ | 0,1 µM y 0,07 µM | En Singleplex las sondas FAM_13104_A y FAM_13124_G | 60°C durante 30'' |
| 4 | Muestras de referencia adicionadas para su sonda específica ³ | 0,07 µM | Duplex las dos parejas de sondas | 60°C durante 30'' |
| 5 | Muestras de referencia adicionadas para su sonda específica ³ | 0,07 µM | Duplex las dos parejas de sondas | 58°C durante 30'' |
| 6 | Muestras de referencia adicionadas para su sonda específica ³ | 0,07 µM | Duplex las dos parejas de sondas | 62°C durante 45'' |
| 7 | Muestras de referencia adicionadas para su sonda específica ³ | 0,4 µM de cebadores y 0,2 µM de sondas | Duplex las dos parejas de sondas | <ul style="list-style-type: none"> • 58°C durante 30'' (sondas 13104) • 60°C durante 30'' (sondas 13124) |

¹Se empleó el mismo pool utilizado en la optimización de la concentración de cebadores.

²Se preparó un pool con las tres muestras sintéticas de referencia en la misma proporción. Para los tres se empleó una muestra diluida que contenía 3,37x10³ copias/µl.

³En singleplex se empleó la muestra de referencia A para las sondas FAM_13104_A y Cy5_13124_A, la muestra de referencia B para la sonda HEX_13104_C y la muestra de referencia E para la sonda FAM_13124_G. En duplex se adicionó en proporciones iguales las muestras de referencia A y B para las sondas FAM_13104_A y HEX_13104_C; y las muestras de referencia A y E para las sondas Cy5_13124_A y FAM_13124_G. En todos los casos se empleó una muestra diluida que contenía 3,37x10³ copias/µl.

*En sombreado se destacan los cambios realizados en las variables para cada una de las pruebas realizadas.

Las reacciones en singleplex y duplex se prepararon respectivamente en base a los volúmenes dispuestos en las tablas 10 y 11, variando el volumen de agua según las diferentes concentraciones de sondas estudiadas. Las reacciones se prepararon y realizaron en las mismas placas y equipos empleados en el resto de experimentos. El programa seleccionado fue el mismo que el representado en la Figura 7, sin embargo, las condiciones de la etapa de annealing/extensión se modificaron de acuerdo a la Tabla 9. Además, no se incluyó la curva de melting, ya que las sondas son específicas de las secuencias de interés y no es necesario analizar la especificidad de la reacción. Por otra parte, también se fue variando el tipo de muestra de ADN empleado, de muestras extraídas en el laboratorio, a las muestras de referencia de ADN sintético. Esto se realizó debido a que la presencia de inhibidores e impurezas en la muestra, condiciona la detección de los niveles de RFU (BIO-RAD Laboratories, 2021). De esta forma, empleando muestras sintéticas más “purificadas”, se quiso descartar que su posible presencia hubiera podido sesgar los resultados obtenidos. Cada muestra se incluyó por triplicado, además del correspondiente blanco.

Tabla 10. Volumen de los reactivos (µl) para la optimización del funcionamiento de las sondas en reacciones en singleplex.

| Reactivos | Volumen (µl) | | |
|------------------------|--------------|------|------|
| SYBR Probe (2x) | 5 | | |
| Primer forward (10 µM) | 0,2 | | |
| Primer reverse (10 µM) | 0,2 | | |
| Sonda x (10 µM) | 0,1 | 0,07 | 0,05 |
| Agua | 2,5 | 2,53 | 2,55 |
| ADN | 2 | | |
| Volumen total (µl) | 10 | | |

Tabla 11. Volumen de los reactivos (μ l) para la optimización del funcionamiento de las sondas en reacciones en duplex.

| Reactivos | FAM_13104_A + HEX_13104_C | | Cy5_13124_A + FAM_13124_G | |
|-----------------------------|---------------------------|------|---------------------------|------|
| SYBR Probe (2x) | 5 | | 5 | |
| Primer forward (10 μ M) | 0,2 | 0,4 | 0,2 | 0,4 |
| Primer reverse (10 μ M) | 0,2 | 0,4 | 0,2 | 0,4 |
| Sonda x + y (10 μ M) | 0,1 | 0,07 | 0,1 | 0,07 |
| Agua | 2,5 | 2,53 | 2,5 | 2,53 |
| ADN | 2 | 2 | 2 | 2 |
| Volumen total (μ l) | 10 | | 10 | |

Las dos variables estudiadas en las diferentes pruebas fueron la concentración adicionada de cada sonda y las condiciones de la etapa de annealing/extensión como posible condicionante del funcionamiento de las sondas.

- Concentración de sondas

En la primera prueba en singleplex se fijaron las concentraciones iniciales de 0,1 μ M y 0,05 μ M, para cada una de las sondas. Estas corresponden a la mitad y un cuarto de la concentración de los cebadores (0,2 μ M cada cebador). La mitad se seleccionó puesto que es una de las proporciones empleadas en trabajos anteriormente publicados sobre la puesta a punto de la qPCR en la discriminación alélica de la kappa-caseína (Kovalchuk et al., 2019; Vafin & Gilmanov, 2021). Por otra parte, puesto que uno de los objetivos es economizar la técnica, también se decidió probar el funcionamiento de las sondas a un cuarto de la concentración de los cebadores. No obstante, en base a los resultados obtenidos y posteriormente detallados, la concentración de 0,05 μ M se consideró insuficiente y se decidió incrementar a 0,07 μ M en las siguientes pruebas.

- Temperatura y duración de la fase de annealing/extensión

Durante la fase annealing/extensión las sondas hibridan específicamente al amplicón, lo cual permite la liberación y detección de la señal fluorescente. Por este motivo, la temperatura y duración de la fase annealing, pueden ser dos grandes condicionantes en el funcionamiento de las sondas. Para fijar las diferentes temperaturas se realizó una búsqueda bibliográfica de estudios con sondas fundamentadas en la misma química, en las que se evaluaba su funcionamiento en diferentes gradientes térmicos. Es preciso señalar que las sondas TaqMan diseñadas por IDT, denominadas “Affinity plus”, son sondas que incorporan 6 monómeros de ácido nucleico bloqueado (LNA). Estos monómeros ofrecen una mayor estabilidad estructural, permitiendo que a pesar de que las sondas tengan una longitud más pequeña para mejorar la discriminación alélica, su T_m sea más elevada. No obstante, en comparación a otras sondas de ADN, la temperatura sigue siendo relativamente menor, entorno a los 55°C, tal y como detallan las guías IDT (BIO-RAD Laboratories, 2006) y se puede observar en la información de la Tabla 4 sobre las sondas diseñadas. Consecuentemente, una menor temperatura melting, también implica una menor temperatura de anillamiento. Según detallan los estudios sobre las sondas LNA, mientras la temperatura supere en 2 o 3°C a la temperatura melting, su eficiencia en la discriminación alélica será óptima (Ugozzoli et al., 2004). De la misma forma, estos coinciden que en comparación a las sondas habituales de ADN, estas son efectivas en un intervalo de temperaturas mucho más amplio (Latorra et al., 2003). Esto último se constató en intervalos de “bajas” (de 58,9°C a 44,3°C) (Ugozzoli et al., 2004) y “altas” temperaturas (de 56.5°C a 68°C) (Brugè et al., 2009). Por todo ello, finalmente se decidió evaluar la aptitud de las sondas en $\pm 2^\circ\text{C}$ a la temperatura ya programada de 60°C. Por otra parte, para terminar de optimizar el funcionamiento de algunas sondas, también se

decidió aumentar la duración de la fase annealing, ya que en las mismas publicaciones se constataron duraciones de 30 y 45 segundos.

Las condiciones de ambas variables se fijaron en el momento en el que los resultados se consideraron óptimos para la prueba realizada. Concretamente, esto significó valores bajos de C_q, niveles de End RFU relativamente altos, y que ambos reflejasen baja variabilidad entre las tres repeticiones incluidas por muestra. Además, también se buscó que las sondas hibridasen correctamente con su secuencia específica y que no se produjera la aparición de dímeros de sondas o la amplificación de otros productos inespecíficos. No obstante, aunque algunos resultados fueron óptimos, se decidió realizar una consulta a la casa comercial que diseñó las sondas. Finalmente, en base a sus recomendaciones y a los resultados obtenidos en las pruebas realizadas hasta el momento, se realizó una última prueba (prueba N° 7) con la que se consiguieron fijar las condiciones ideales para el correcto funcionamiento de las sondas.

3.5. Validación del método de identificación de alelos A, B, y E del gen *CSN3* mediante qPCR

Para poder asegurar el correcto funcionamiento de la técnica desarrollada en base a las condiciones fijadas, se debe realizar la validación del ensayo. Para ello, existen diferentes guías de referencia que definen determinados parámetros, en base a los que se debe evaluar una técnica. En este caso, se analizó la eficiencia, sensibilidad, rango dinámico y la repetibilidad, de acuerdo a lo establecido en las guías publicadas por los organismos Codex Alimentarius (2010) y Food and Drug Administration (FDA) (2019), para la validación de PCRs cuantitativas.

3.5.1. Rango dinámico, límite de detección (LOD) y cuantificación (LOQ)

El rango dinámico se define como el rango de concentraciones en el que el fragmento de interés se detecta y cuantifica de forma fiable. De esta forma, este también permite definir las curvas estándar que pueden emplearse para realizar cuantificaciones absolutas. Este rango se establece conforme a dos valores, el límite de detección (Limit of Detection o LOD), que también estima la sensibilidad del ensayo, y el límite de cuantificación (Limit of Quantification o LOQ). El primero se define como la concentración más baja a la que se detecta el 95% de las muestras positivas, y el segundo, corresponde a la cantidad más baja de ADN en una muestra que puede cuantificarse de forma fiable. Una forma de determinar el valor LOQ, es determinar la cantidad mínima de ADN diana, con la que las réplicas realizadas dan un resultado positivo con un coeficiente de variabilidad (CV) del C_q no superior a 0,5 C_q. Para validar un rango dinámico en el ensayo se deben emplear tres réplicas por muestra y el valor R² de las curvas estándar debe ser $\geq 0,98$. El rango dinámico se considera óptimo si presenta como mínimo un rango lineal de 5 órdenes de magnitud.

3.5.2. Eficiencia de la amplificación

La eficiencia de la amplificación también se comprueba mediante las curvas estándar. Con estas se determina el valor C_q, de cada dilución que las conforma, para posteriormente trazar el logaritmo de la cantidad de ADN frente a su valor C_q. De esta forma, con la pendiente media de la curva estándar, se puede calcular la eficiencia del ensayo. La pendiente de la curva estándar se estima mediante la ecuación de la recta $y = mx + b$, donde $y = C_q$ y $x = \log_{10}$ de la cantidad de ADN. Posteriormente, la eficiencia de la qPCR, se calcula mediante la ecuación, $10^{-1/\text{pendiente}} - 1$. Para que un ensayo se considere óptimo, la pendiente debe encontrarse entre los valores de -3,1 y -3,6, lo cual equivale a una eficiencia del 90%-110%.

3.5.3. Repetibilidad, intra-assay e inter-assay repeatability

Por último, la precisión de un ensayo, se puede evaluar mediante la repetibilidad intra-assay e inter-assay. En este caso, al igual que en el trabajo publicado por Peixoto y Marcelino (2012), la repetibilidad intra-assay se evaluó por medio del coeficiente de variación (CV) de los valores Cq de las repeticiones introducidas en una misma placa por muestra. En cuanto a la repetibilidad inter-assay, esta también se evaluó por medio del CV de los resultados obtenidos mediante la repetición de cada placa tres veces (Peixoto et al., 2005).

En definitiva, para validar la técnica con respecto a estos tres parámetros, primero se prepararon tres curvas estándar, una para cada variante alélica estudiada. Una curva estándar es una serie de diluciones preparadas a partir de una muestra de referencia de concentración o número de copias conocida. De esta forma, se prepararon siete diluciones seriadas a partir de las tres muestras de referencia sintéticas, para las que se calculó un número de copias de $6,7 \times 10^{10}$ copias/ μ l (Anexo 2). Puesto que este se consideró un número elevado, se decidió diluir las muestras a $6,7 \times 10^6$ copias/ μ l. Finalmente, se realizaron diluciones seriadas 10-fold desde las $6,7 \times 10^6$ copias/ μ l hasta las $6,7 \times 10^0$ copias/ μ l. Posteriormente, estas se volvieron a diluir a la mitad, ya que en las reacciones se adicionan 2 μ l de muestra de ADN. De esta forma, se añadiría el número de copias deseado.

Para preparar las reacciones en duplex, se emplearon los volúmenes fijados tras optimizar la técnica, aquellos que fueron establecidos para conseguir una concentración final de 0,4 μ M de cada cebador y de 0,2 μ M para cada sonda de cada pareja (Tabla 11). Además, para calcular el rango dinámico y evaluar la repetibilidad conforme a lo definido previamente, se incluyeron tres réplicas por muestra. También se incluyeron las tres muestras correspondientes al blanco, y cada placa se repitió tres veces. Por último, puesto que la finalidad de la técnica es lograr la discriminación alélica, cada muestra de referencia se amplificó con cada pareja de sondas, seleccionando en el equipo los programas definidos tras la optimización de la técnica (Tabla 16). Esto también permitió evaluar la especificidad de cada sonda.

Una vez obtenidos los resultados, primero se estimó el rango dinámico de las tres placas realizadas para cada alelo y para cada pareja de sondas. Para ello, se ajustó su límite inferior, de tal forma que se cumplieran los requisitos para los parámetros LOD y LOQ. De esta forma, se estableció como límite aquella dilución en la que el fragmento amplificado se detectase con una probabilidad superior al 95% y cuyo resultado positivo tuviera un valor CV inferior a 0,5. Además, también se quiso asegurar la especificidad de la amplificación, por lo que se descartaron aquellas diluciones en las que se detectase un falso positivo, es decir, las diluciones en las que la sonda no específica para la variante analizada emitiera fluorescencia. Esto también se corroboró mediante los gráficos de discriminación alélica trazados por el software del equipo. Por otra parte, para descartar la posible dimerización de sondas u otras amplificaciones inespecíficas, también se descartaron aquellas diluciones en las que el Cq medio de las muestras, fuese superior al Cq medio del control negativo o blanco. A continuación, se estimó la pendiente, eficiencia y valor de R^2 de las curvas ya acotadas, para asegurar su eficiencia. Por último, se calculó la repetibilidad intra- e inter-assay.

3.6. Aplicabilidad de la qPCR en la identificación de los alelos A, B y E de la kappa-caseína

Finalmente, para corroborar la consecución de los objetivos fijados en el presente trabajo, con el ensayo seguidamente descrito, se pretendió evaluar la aplicabilidad de la técnica desarrollada en el genotipado del ganado vacuno de raza Frisona. Para ello, se emplearon las muestras del ganado genotipado, muestras de leche y pelo procedentes del mismo animal. De esta forma, se escogió una muestra representativa para cada uno de los seis posibles genotipos. Los volúmenes y condiciones bajo los que se llevaron a cabo las reacciones, fueron los definidos en la técnica

optimizada (Tabla 16). Por último, la conformidad de los resultados se evaluó a partir de los datos obtenidos de las curvas de amplificación y los gráficos de discriminación alélica.

4. RESULTADOS

4.1. Extracción de ADN a partir de células somáticas de leche y de pelo

Los resultados del ANOVA simple realizado para seleccionar la matriz que ofrece mejores resultados en la extracción, quedan representados en la Tabla 12. Este análisis descompone la variabilidad en dos fuentes, permitiendo comprobar si existen más diferencias entre los diferentes grupos muestrales de la extracción (variabilidad entre grupos) o dentro de los mismos (variabilidad intra grupos), para las diferentes variables definidas. De esta forma, permite comprobar si el efecto general de la extracción es significativo para cada una de las variables. Puesto que el nivel de significación fijado fue del 95%, se puede asumir que existen diferencias significativas con un 5% de error, si la probabilidad o valor-P de la prueba-F es inferior a 0,05.

Tabla 12. Análisis de la varianza del factor “extracción”.

| Variables | F | P-valor* |
|-----------------------------|-------|--------------|
| Concentración (ng/μl) | 4,25 | 0,009 |
| A _{260/280} (nm) | 34,47 | 0,000 |
| A _{260/230} (nm) | 18,29 | 0,000 |
| <i>Amplificability</i> (Cq) | 19,32 | 0,000 |

*Se destaca en sombreado si existen diferencias significativas ($P \leq 0,05$).

Los resultados reflejan que el factor “extracción”, presenta efectos significativos en las cuatro variables estudiadas. No obstante, este análisis no permite conocer la significancia del efecto individual de cada uno de los tipos de muestras empleados en la extracción. Por ello, a continuación, se realiza una prueba multivariante que ofrece resultados para cada tipo de muestra y el contraste entre las mismas. En este caso, las diferencias son significativas si la distancia entre las medias es igual o superior a un límite establecido por el método empleado (en este caso el método Tukey HSD). Esto a su vez depende del error tipo I (alfa) asumido, que es del 5%. En definitiva, los resultados de esta prueba, permiten diferenciar para cada una de las variables estudiadas, los tipos de muestras que presentan diferencias significativas de las que no. Además, estos también permiten establecer subconjuntos homogéneos en los que se agrupan las muestras estadísticamente iguales (Tabla 13).

Tabla 13. Efecto de la temperatura de conservación de las muestras y tipo de matriz en la cantidad (ng/μl), calidad (A_{260/280} y A_{260/230}) y *amplificability* (Cq) de las muestras de ADN.

| Muestras empleadas en la extracción | Concentración ¹ (ng/μl) | | A _{260/280} ¹ (nm) | | A _{260/230} ¹ (nm) | | <i>Amplificability</i> ² (Cq) | |
|-------------------------------------|---------------------------------------|--------|---|-------|---|-------|---|-------|
| | Media | SD* | Media | SD* | Media | SD* | Media | SD* |
| Leche a 4°C | 13,07 ^{ab} | 11,274 | 1,36 ^c | 0,221 | 0,40 ^c | 0,173 | 24,62 ^a | 1,96 |
| Leche a -20°C | 9,69 ^b | 10,784 | 1,71 ^b | 0,185 | 2,76 ^a | 1,669 | 23,30 ^a | 0,863 |
| Leche a -80°C | 20,99 ^a | 2,652 | 1,92 ^a | 0,048 | 2,22 ^{ab} | 0,249 | 18,15 ^b | 5,231 |
| Pelo | 10,74 ^b | 9,453 | 1,93 ^a | 0,168 | 1,38 ^b | 0,523 | 26,29 ^a | 1,035 |

^{abc}Los valores con el mismo subíndice son estadísticamente iguales entre sí ($P > 0,05$) y presentan diferencias significativas con los valores de subíndices diferentes ($P \leq 0,05$).

¹Valores más elevados indican mejores resultados.

²Valores más bajos indican mejores resultados.

*SD = Standard Deviation.

Los datos recopilados reflejan lo dispuesto a continuación:

- **Concentración de ADN:** La cantidad de ADN extraído de las muestras de leche conservada a -80°C fue significativamente mayor al resto de muestras ($P \leq 0,05$), pero igual a la cantidad extraída de las muestras de leche refrigerada ($P > 0,05$). No obstante, la cantidad de ADN extraído a partir de leche refrigerada, reflejó una variabilidad alta (valores elevados para la desviación típica o SD) que pudo sesgar el valor de la media. Además, la leche refrigerada no presenta diferencias significativas con el resto de muestras ($P > 0,05$).
- **Valores de $A_{260/280}$:** La calidad de ADN procedente de las muestras de pelo y leche conservada a -80°C fue estadísticamente mejor que la calidad del ADN extraído a partir de la leche conservada a -20°C y 4°C (ambos reflejaron valores comprendidos entre 1,8 y 2).
- **Valores de $A_{260/230}$:** La calidad del ADN extraído de leche conservada a -20°C y a -80°C , fue estadísticamente igual y la mejor, con respecto al resto de subconjuntos (los valores se aproximan más a 2,2). No obstante, para las muestras del pelo no se encontraron diferencias significativas con las muestras de leche conservada a -80°C , pero sí con las muestras de leche a -20°C y 4°C . Por otra parte, es notorio que la calidad del ADN extraído de la leche a -20°C , presenta una gran variabilidad, que también ha podido incrementar el valor de la media.
- **Amplificability:** En este caso, aunque con una mayor variabilidad, con la leche congelada a -80°C se extrajo una mayor cantidad de ADN (en el resto de muestras no se evidenciaron diferencias significativas ($P > 0,05$)). Puesto que este resultado guarda correlación con la concentración de ADN ($\text{ng}/\mu\text{l}$) cuantificada, se puede asumir que la calidad del ADN extraído no ha comprometido su amplificación por qPCR.

Finalmente, a rasgos generales, se puede considerar que la cantidad y calidad de ADN extraído fue mejor y menos variable en el mismo orden: leche congelada a -80°C , pelo, leche conservada a -20°C y leche refrigerada. Por otra parte, se ha observado que la temperatura a la que se conservan las muestras genera diferencias significativas en el ADN extraído, por lo que es preferible que, en caso de extraer ADN a partir de leche, las muestras se conserven previamente a -80°C . Este efecto de la temperatura, junto a la dificultad que supone la extracción de ADN a partir de leche dada la elevada presencia de sustancias inhibitoras y la gran variabilidad del contenido de células, también fue evidenciado en el trabajo realizado por Wassermann (2020). El mismo autor asoció el fenómeno de las bajas temperaturas al daño mecánico ocasionado en los tejidos y paredes celulares, debido a la formación de cristales de hielo durante el proceso de congelación y posterior descongelación. Esto favorece a una mayor liberación de ADN y a la posterior digestión con la proteinasa K, lo cual posibilita una extracción más eficaz (Wassermann, 2020).

Por último, es oportuno considerar que los resultados del pelo podrían haberse visto afectados negativamente por la pequeña y variable cantidad empleada durante las extracciones. Es preciso recordar que la extracción a partir de leche ya había sido previamente optimizada en el laboratorio, por lo que siempre se empleó la misma cantidad de muestra. Por este motivo, un estudio para optimizar la extracción de ADN a partir de pelo podría ofrecer mejores resultados, y por ello, no se descarta como posible matriz para la extracción.

Para finalizar en lo que concierne a la extracción, es preciso señalar que para lograr resultados representativos en la prueba de genotipado (descrita en el apartado 3.6 de la metodología), se seleccionaron las muestras que reflejaron los mejores resultados en la extracción. Para ello, se

realizó una comparativa entre las muestras de pelo y leche procedentes del mismo individuo. Dichos resultados figuran en el Anexo 3, junto con las muestras finalmente seleccionadas.

4.2. Optimización de las condiciones de amplificación mediante qPCR

4.2.1. Optimización de la concentración de cebadores

De forma previa a la interpretación de los resultados, para asegurar que los datos en los que se basó el análisis estadístico eran óptimos (Tablas 14 y 15), se evaluó la curva de disociación del ensayo realizado (Figura 8). Con esta se pudo comprobar que en todas las muestras amplificadas sólo amplificó el fragmento de interés del gen *CSN3*. De la misma forma, se corroboró la especificidad de la reacción y el correcto funcionamiento de los cebadores diseñados, puesto que no amplificaron otros productos y se descartó la posible formación de dímeros de cebadores.

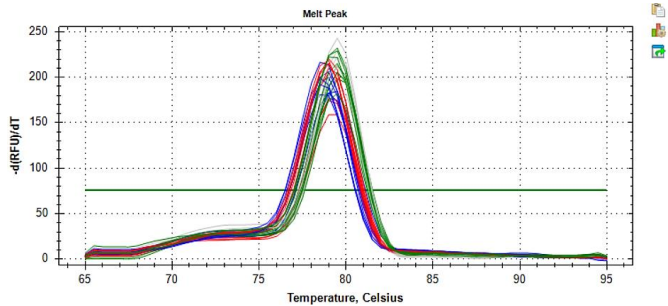


Figura 8. Curva de disociación de la qPCR. Cada curva representa una de las muestras de la matriz experimental diseñada (en azul las muestras con 0,2 μ M de cebadores, en rojo con 0,3 μ M de cebadores y en verde con 0,4 μ M de cebadores).

Tabla 14. Análisis de la varianza del factor “extracción”.

| Variables | F | P-valor* |
|-----------|--------|--------------|
| Cq | 3,020 | 0,024 |
| End RFU | 13,064 | 0,000 |

*Se destaca en sombreado si existen diferencias significativas ($P \leq 0,05$).

Tabla 15. Efecto de la concentración de cebadores y cantidad de muestra en valores Cq y niveles de fluorescencia alcanzados al final de la amplificación (End RFU) de las muestras de ADN.

| Cantidad de ADN | Concentración de cebadores | Cq ¹ | | End RFU ² | |
|-----------------|----------------------------|--------------------|-------|-----------------------|--------|
| | | Media | SD* | Media | SD* |
| 30 ng | 0,2 μ M | 19,93 ^a | 0,250 | 788,36 ^a | 61,495 |
| | 0,3 μ M | 19,67 ^a | 0,040 | 772,2 ^a | 29,516 |
| | 0,4 μ M | 19,48 ^a | 0,121 | 748,84 ^{ab} | 23,125 |
| 45 ng | 0,2 μ M | 20,06 ^a | 0,051 | 680,79 ^{abc} | 7,782 |
| | 0,3 μ M | 19,57 ^a | 0,144 | 686,64 ^{abc} | 13,989 |
| | 0,4 μ M | 19,94 ^a | 0,222 | 545,41 ^d | 54,328 |
| 60 ng | 0,2 μ M | 19,63 ^a | 0,275 | 625,38 ^{bcd} | 77,747 |
| | 0,3 μ M | 19,49 ^a | 0,282 | 597,59 ^{cd} | 63,595 |
| | 0,4 μ M | 19,92 ^a | 0,352 | 528,15 ^d | 22,937 |

^{abcd} Los valores con el mismo subíndice son estadísticamente iguales entre sí ($P > 0,05$) y presentan diferencias significativas con los valores de subíndices diferentes ($P \leq 0,05$).

¹Valores más bajos indican mejores resultados.

²Valores más elevados indican mejores resultados.

*SD = Standard Deviation.

El análisis estadístico indica que, para ambas variables fijadas, existen diferencias significativas entre las diferentes combinaciones de cebadores y ADN (Tabla 14). En el caso de los valores de

Cq, pese a que el ANOVA simple sí detecta diferencias significativas, posteriormente la prueba post-hoc considera que las diferentes combinaciones amplifican el gen en un número de ciclos estadísticamente iguales (Tabla 15). Esto se debe a que el método empleado impone condiciones más exigentes en la detección de diferencias significativas que otros métodos como el método Duncan, con el que sí se llega a diferenciar diferentes subconjuntos. No obstante, puesto que en el ámbito práctico la variabilidad de los resultados está sujeta a muchos factores, errores instrumentales, el factor humano... se considera que las diferencias detectadas con métodos menos restrictivos, podrían resultar en falsos positivos. Por este motivo, se decide mantener el método Tukey HSD hasta ahora aplicado, y basar la elección en la otra variable estudiada para el que sí se detectan diferentes subconjuntos.

Aunque los niveles de las unidades relativas de fluorescencia o RFU, no condicionan el valor obtenido de Cq, sí es cierto que mayores valores favorecen a una mejor visualización de los resultados (BIO-RAD Laboratories, 2021). Un mayor valor de RFU al final de la reacción sugiere que la fase meseta o “plateau” de la curva de amplificación alcanzada es más elevada, y por ello, que la cantidad de producto amplificado es mayor. De esta forma, se comprueba que independientemente de la cantidad inicial de ADN presente en la muestra, los cebadores en una concentración final de 0,2 μM o 0,3 μM , amplifican cantidades significativamente iguales ($P>0,05$) (Tabla 15). Además, estas son algo más eficientes que la concentración final de 0,4 μM , ya que en los casos en los que se emplearon 45 ng de ADN, los valores alcanzados de “End RFU” con 0,2 μM y 0,3 μM fueron significativamente mayores ($P\leq 0,05$). Por otra parte, también se observa que los niveles más elevados de fluorescencia se obtuvieron con 30 ng de ADN, por lo que una menor cantidad de muestra habría resultado en una mejor amplificación.

Finalmente, puesto que uno de los objetivos en el desarrollo de la técnica es la economización de la misma, para discernir entre las dos mejores concentraciones, se seleccionó aquella en la que se precisa de una menor concentración de cebadores. En definitiva, esto supone que para lograr la correcta amplificación del gen *CSN3*, cada uno de los cebadores de la pareja de primers diseñados deben encontrarse en la reacción en una concentración de 0,2 μM .

4.2.2.Optimización del funcionamiento de las sondas

En el Anexo 1, a modo de resumen, se esquematizan las condiciones y resultados de las pruebas que fueron realizadas hasta mejorar la eficiencia de las dos parejas de sondas en reacciones duplex. Por medio de las primeras pruebas realizadas en singleplex, se estandarizó la concentración de las sondas. Posteriormente, con las pruebas en duplex, se fijaron las condiciones de la etapa annealing.

- Concentración de sondas

Los resultados de las primeras pruebas realizadas en singleplex reflejaron un funcionamiento muy diferente en la emisión de la señal fluorescente entre las cuatro sondas. Mientras que las sondas marcadas por los fluorocromos Cy5 y HEX emitieron valores más elevados de RFU, las sondas marcadas con FAM reflejaron valores muy bajos y variables. Por consiguiente, para seleccionar la concentración más óptima de las sondas, se emplearon los resultados correspondientes a las sondas marcadas con los fluoróforos Cy5 y HEX, ya que fueron más representativos. Para ello, se evaluó el efecto de la concentración en las variables “Cq” y “RFU” (ANOVA simple), cuyos resultados también se resumen en el Anexo 1.

Con dicho análisis estadístico, se constató que no existen diferencias significativas entre los valores Cq obtenidos con las dos concentraciones estudiadas ($P>0,05$). De esta forma, para escoger una de las concentraciones, la selección se basó en el valor de RFU para el que sí se

observaron diferencias significativas ($P \leq 0,05$), aunque cada sonda reflejó mayores valores de RFU para una concentración diferente, la sonda HEX_13104_C para una concentración de sonda de $0,07 \mu\text{M}$ y la sonda Cy5_13124_A para $0,1 \mu\text{M}$. Puesto que también se evidenció que los niveles de RFU entre las sondas Cy5_13124_A y HEX_13104_C, eran más próximos empleando una concentración de $0,07 \mu\text{M}$, finalmente se seleccionó esta concentración como óptima. Estos valores de RFU más balanceados permiten una mejor discriminación alélica mediante los gráficos que genera el software del termociclador (Figuras 9 y 10). Adicionalmente, esta concentración también permite economizar la técnica, ya que supone el empleo de una menor cantidad de sondas.

- Temperatura y duración de la fase de annealing/extensión

Tras la primera prueba de la qPCR en duplex, en la que se evaluó el funcionamiento de las parejas de sondas en una misma reacción, se comprobó que los resultados de las sondas FAM continuaron siendo insatisfactorios. De esta forma, el resto de pruebas se enfocó en la variación de las condiciones de temperatura y tiempo de la etapa annealing. Finalmente, se observó que la pareja de sondas que discrimina las variantes con mutación en la posición 13104, logra una mejor hibridación y emisión de fluorescencia con una temperatura annealing de 58°C . Por el contrario, las pruebas a diferentes tiempos y temperatura no consiguieron mejorar el funcionamiento de la sonda FAM_13124_G, y el funcionamiento con respecto a su pareja (Cy5_13124_A) continuó siendo desbalanceado.

En vista a la imposibilidad de mejorar el funcionamiento de la sonda FAM_13124_G, se decidió realizar una consulta a la casa comercial IDT. Tras evaluar los resultados, se consideró que la gran discrepancia en el funcionamiento de las sondas, pudo deberse a la complejidad de su diseño. De esta forma, aunque algunas sondas sí funcionaron bajo las condiciones inicialmente estudiadas, para lograr maximizar y equiparar el rendimiento de todas ellas, se recomendó incrementar las concentraciones de cebadores a $0,4 \mu\text{M}$ y de sondas a $0,2 \mu\text{M}$. No obstante, puesto que disminuir la temperatura annealing logró minimizar las diferencias entre los niveles de RFU de las sondas de la posición 13104, esta condición sí se mantuvo.

En la Tabla 16 se recogen los resultados obtenidos en la última prueba realizada, con la que finalmente se consideró optimizado el funcionamiento de las sondas en duplex para la discriminación alélica. Tal y como se comprueba, los valores medios de C_q obtenidos con las cuatro sondas, son bajos y relativamente similares. Teniendo en cuenta que las muestras empleadas contenían la misma cantidad de ADN, esto denota un funcionamiento óptimo. No obstante, es preciso señalar que comparativamente, la sonda FAM_13124_G fue detectada en ciclos más tardíos. En cualquier caso, puesto que la técnica tiene como prioridad discriminar las variantes alélicas y no cuantificar las cantidades de ADN presentes en la muestra, las pequeñas diferencias entre los valores de C_q de la sonda no guardan gran relevancia.

Tabla 16. Resultados de la prueba N°7 en las que se fijan las condiciones de las cuatro sondas para la detección de las variantes A, B y E de la kappa-caseína.

| Sondas | Cq ¹ | | End RFU ² | | Condiciones fijadas | | |
|-------------|-----------------|-------|----------------------|--------|-------------------------|----------------------|--------------------------|
| | Media | SD* | Media | SD* | Concentración Cebadores | Concentración Sondas | Tª y tiempo annealing |
| FAM_13104_A | 27,12 | 0,074 | 192,53 | 5,442 | 0,4 μM | 0,2 μM | 58°C durante 30 segundos |
| HEX_13104_C | 26,92 | 0,110 | 405,28 | 19,494 | | | |
| Cy5_13124_A | 24,94 | 0,032 | 693,00 | 27,483 | | | 60°C durante 30 segundos |
| FAM_13124_G | 31,83 | 1,596 | 65,39 | 7,476 | | | |

¹Valores más bajos indican mejores resultados.

²Valores más elevados indican mejores resultados.

*SD = Standard Deviation.

Con respecto a los niveles de RFU, aunque consiguieron aumentarse con respecto a las anteriores pruebas, los niveles de las sondas marcadas con el fluoróforo FAM continúan reflejando valores inferiores. Consecuentemente, se comprueba que existen diferencias entre las parejas de sondas, siendo más notorias, entre las sondas de la posición 13124. Además, en comparación con la discriminación alélica a partir de los valores C_q , este desequilibrio sí podría dificultar la identificación de las variantes si se realiza por medio de los gráficos de discriminación alélica. No obstante, esto no pudo comprobarse hasta la evaluación de la aplicabilidad de la técnica.

4.3. Validación del método de identificación de alelos A, B y E mediante qPCR

Para estimar los valores de las pendientes, eficiencias y valor de R^2 , conforme a lo descrito en el apartado 3.5 de la metodología, se hizo uso de los datos proporcionados por las curvas de amplificación y gráficos de discriminación alélica trazados por el software del equipo. A modo de ejemplo, en las Figuras 9 y 10, se muestran los gráficos correspondientes a una de las placas o ensayos realizados para cada pareja de sondas. A partir de los mismos, se realizó el ajuste del rango dinámico de las curvas estándar de cada alelo y cada ensayo. Para ello, se estimaron los valores pertinentes de LOD y LOQ y se aseguró la especificidad y correcto funcionamiento de las sondas para la correcta discriminación alélica.

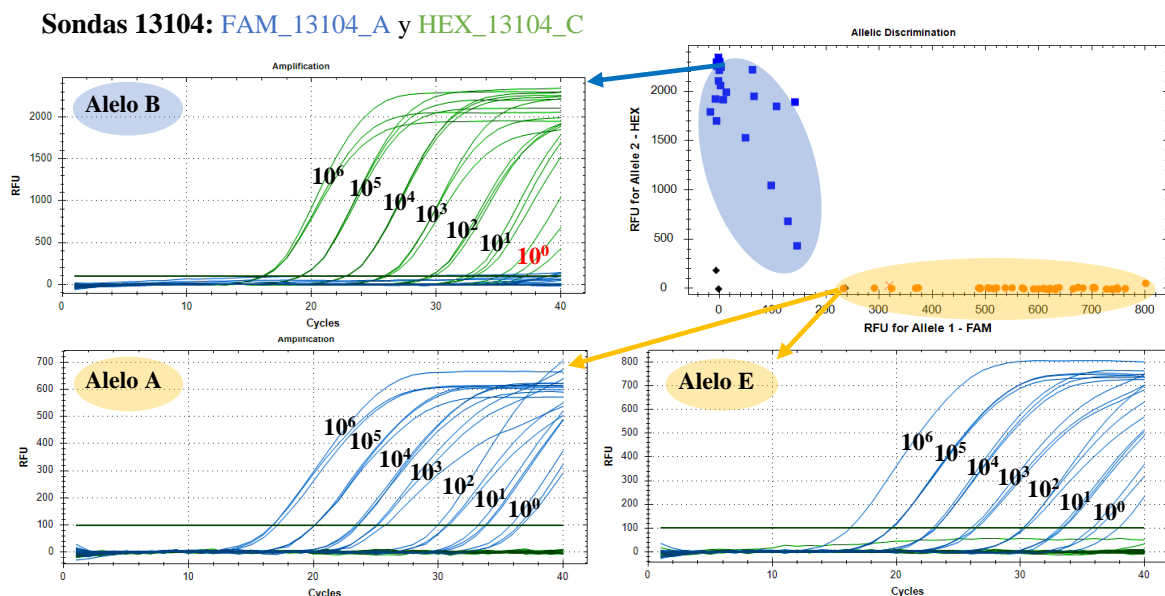


Figura 9. Curvas de amplificación (RFU frente al número de ciclos) de las curvas estándar y gráfico de discriminación alélica (RFU del fluoróforo HEX frente al RFU del fluoróforo FAM) del segundo ensayo. La sonda FAM_13104_A emite la señal azul y es específica de los alelos A y E, mientras que la sonda HEX_13104_C emite la señal verde y es específica de la variante B. Esta pareja permite diferenciar los alelos A y E (“○”), del alelo B (“■”). En este caso, se observa una buena discriminación, ya que en ninguno de los alelos se detecta señal fluorescente de la sonda no específica (se garantiza la especificidad de las sondas). Las diluciones descartadas se encuentran marcadas en rojo y se excluyen para obtener valores óptimos de LOD y LOQ. Los “◆” del gráfico de discriminación no se asocian a ningún alelo y corresponden a los blancos.

Sondas 13124: Cy5_13124_A y FAM_13124_G

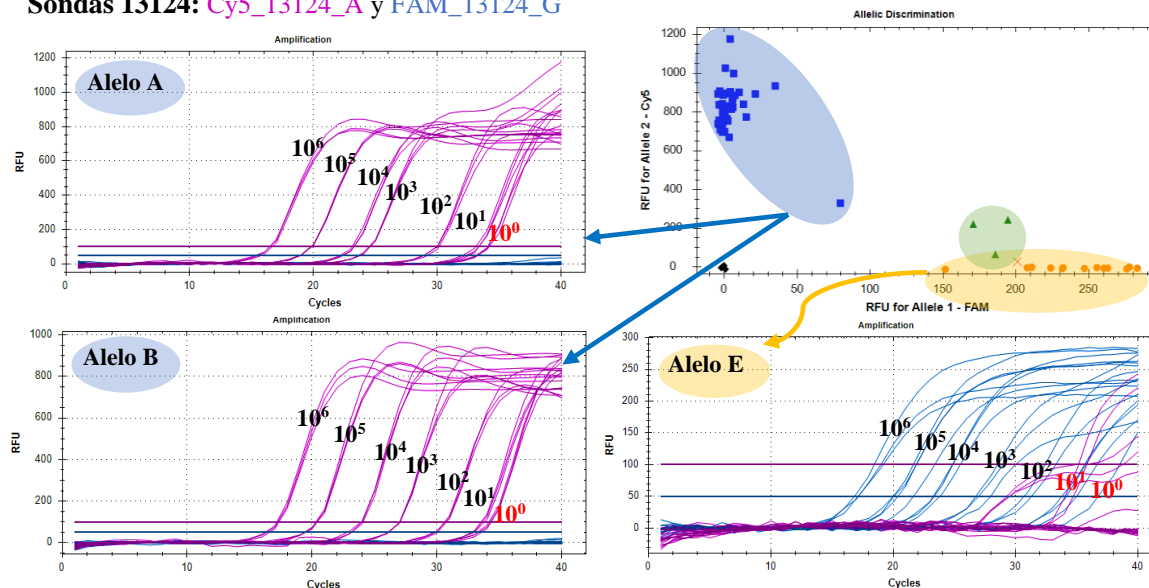


Figura 10. Curvas de amplificación (RFU frente al número de ciclos) de las curvas estándar y gráfico de discriminación alélica (RFU del fluoróforo Cy5 frente al RFU del fluoróforo FAM) del tercer ensayo. La sonda Cy5_13124_A es específica para los alelos A y B y emite la señal morada, y la sonda FAM_13124_G es específica para el alelo E y emite la señal azul. Esta pareja permite discriminar la variante E (“●”) de las otras dos variantes (“◆”). En este caso, la discriminación alélica no es tan precisa puesto que en las curvas de las diluciones 10^1 y 10^0 del alelo E, el equipo detecta las señales de ambas sondas, interpretando el alelo como heterocigoto (“▲”). No obstante, en estas diluciones la interpretación de la señal Cy5 como amplificación positiva, es un falso positivo. Un resultado positivo habría reflejado niveles de RFU mucho mayores (800 - 1000), como los alcanzados con los alelos A y B. Por consiguiente, además de por aspectos experimentales, este error puede deberse a la baja concentración de la muestra y a la gran divergencia en la intensidad de la señal fluorescente entre sondas. Por este motivo, la discriminación alélica no es correcta y ambas diluciones se descartan del rango dinámico. Las diluciones descartadas, marcadas en rojo, también se excluyen para obtener valores óptimos de LOD y LOQ. Los “◆” del gráfico de discriminación no se asocian a ningún alelo y corresponden a los blancos.

Una vez ajustados los rangos dinámicos, se trazaron los gráficos representados en la Figura 11, para obtener las ecuaciones de las rectas que definen la relación del logaritmo de la cantidad de ADN frente al valor Cq. Finalmente, con estas ecuaciones se determinaron y calcularon los tres parámetros descritos en las figuras, también recogidos de forma más esquemática en el Anexo 2.

Tal y como reflejan los resultados obtenidos, las curvas estándar preparadas se ajustan a los criterios de calidad perseguidos, dado que todas presentaron valores de R^2 iguales o superiores a 0,98. Esto permite garantizar que los datos empleados en el cálculo del resto de parámetros, son óptimos y permiten validar la técnica desarrollada. En este caso, puesto que los valores de la eficiencia, también se encuentran dentro del intervalo óptimo (90-110%), se confirma la validez de los rangos dinámicos establecidos en cada ensayo. Por otra parte, es preciso destacar que en algunos ensayos, la eficiencia o calidad de las curvas no se ajustó a lo requerido. No obstante, esto se asocia a errores experimentales puntuales, ya que no se repitió en las respectivas placas homólogas. Por otra parte, se observó que en todos los ensayos realizados, el límite de detección (LOD) pudo fijarse en $6,73 \times 10^0$ copias/ μ l. No obstante, el límite de cuantificación (LOQ) fue más variable, por lo que para unificar los resultados obtenidos y definir el rango dinámico más adecuado para la detección y cuantificación de cada alelo que pudiera estar presente en una muestra, se seleccionó como óptimo el rango más restrictivo. De esta forma, se pretende garantizar la validez del rango definido en futuros ensayos. Además, es preciso destacar que en todos los casos, se logró fijar un rango dinámico que abarcara como mínimo 5 órdenes de magnitud.

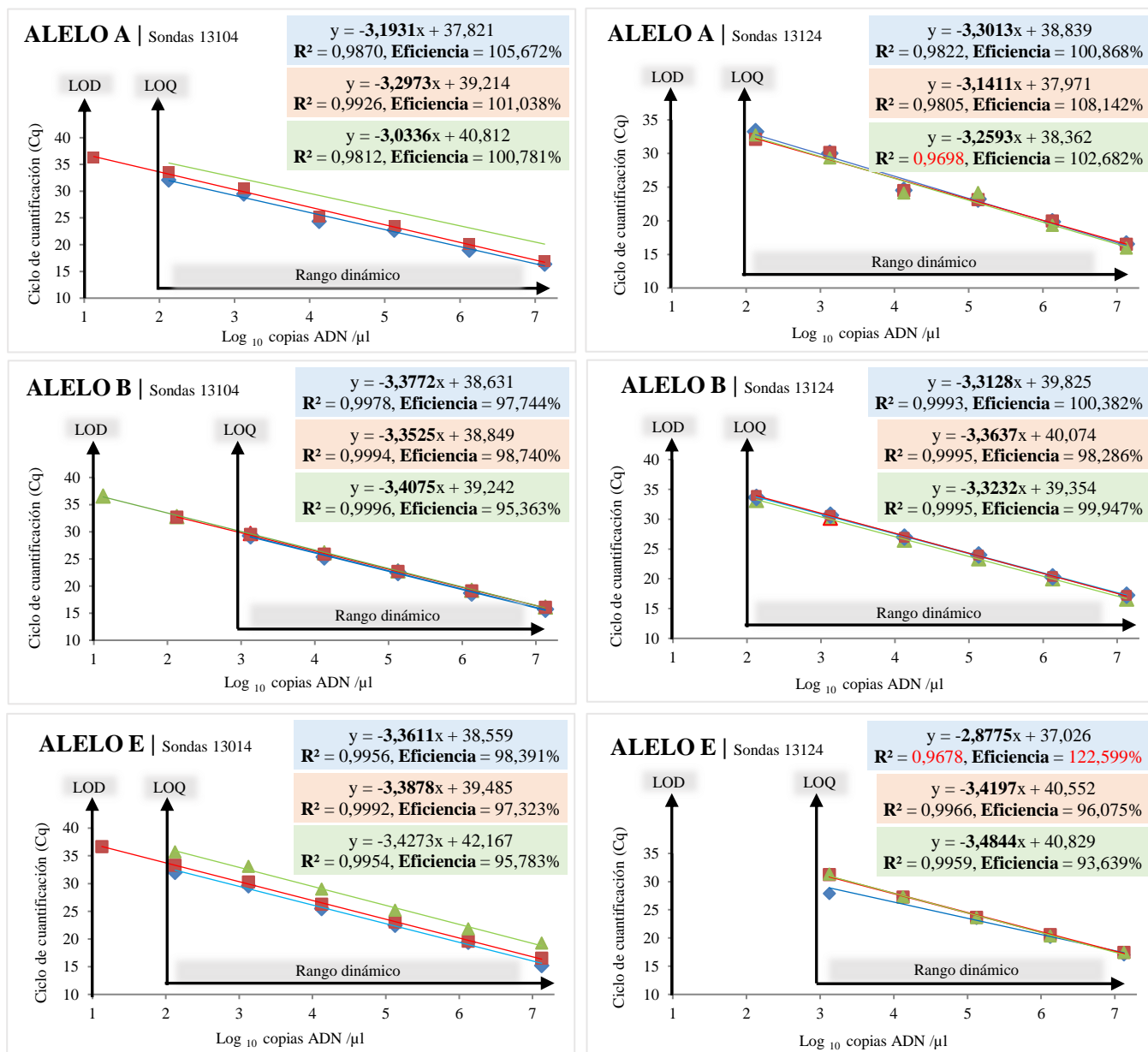


Figura 11. Curvas estándar, con el rango dinámico ajustado, de los tres ensayos realizados para la pareja de sondas de la posición 13104: FAM_13104_A y HEX_13104_C (izquierda) y la pareja de sondas de la posición 13124: Cy5_13124_A y FAM_13124_G (derecha). Con las ecuaciones representadas se estima la pendiente, eficiencia y valor de R^2 ; del primer (azul), segundo (rojo) y tercer (verde) ensayo. En rojo se destacan los valores que no se ajustan a los valores óptimos. Por último, para cada alelo y pareja de sondas, se indican los valores de LOD, LOQ y el rango dinámico más restrictivo de los tres ensayos.

Por último, los resultados de los parámetros calculados para evaluar la repetibilidad intra- e inter-assay de los rangos dinámicos definidos, se recogen en el Anexo 2. En las Tablas 17 y 18, se establecen los valores máximos y mínimos entre los que comprendieron los valores del coeficiente de variación (%). Puesto que en todos los casos, estos fueron inferiores al 25%, se puede asumir la buena repetibilidad de la técnica desarrollada (Codex Alimentarius, 2010; FDA, 2019).

Tabla 17. Intervalos del Coeficiente de Variación (%) de la repetibilidad intra- e inter-assay, entre los que se encuentran los valores del rango dinámico fijado para los tres alelos y la pareja de sondas de la posición 13104 (FAM_13104_A y HEX_13104_C).

| CV (%) | Intra-assay repeatability | | | Inter-assay repeatability | | |
|--------|---------------------------|-------------|-------------|---------------------------|-------------|--------------|
| | Alelo A | Alelo B | Alelo E | Alelo A | Alelo B | Alelo E |
| | 0,09 - 4,60 | 0,12 - 1,19 | 0,22 - 5,47 | 5,35 - 12,87 | 0,66 - 1,45 | 5,59 - 12,13 |

Tabla 18. Intervalos del Coeficiente de Variación (%) de la repetibilidad intra- e inter-assay, entre los que se encuentran los valores del rango dinámico fijado para los tres alelos y la pareja de sondas de la posición 13124 (Cy5_13124_A y FAM_13124_G).

| CV (%) | Intra-assay repeatability | | | Inter-assay repeatability | | |
|--------|---------------------------|-------------|--------------|---------------------------|-------------|-------------|
| | Alelo A | Alelo B | Alelo E | Alelo A | Alelo B | Alelo E |
| | 0,07 - 10,83 | 0,06 - 0,74 | 0,01 - 15,76 | 0,94 - 2,53 | 0,90 - 2,06 | 0,17 - 6,59 |

4.4. Aplicabilidad de la qPCR en la identificación de los alelos A, B y E de la kappa-caseína

Los resultados del genotipado de animales realizado mediante la técnica desarrollada en el presente trabajo se muestran en la Tabla 19. Con esta se deduce que en ambas reacciones duplex las parejas de sondas identificaron correctamente a su variante o sus variantes específicas. En todos los casos las amplificaciones positivas reflejaron valores Cq adecuados y con baja variabilidad entre muestras. No obstante, pese a que en todos los casos las señales se ajustaron a lo esperado, en el caso de las sondas de la posición 13124, en el genotipo BB se detectó fluorescencia de la sonda no específica FAM_13124_G. Esta, además, generó una peor amplificación en los genotipos que contienen su variante específica (alelo E; genotipos AE, BE y EE), ya que las desviaciones estándar de los valores Cq fueron notoriamente más elevadas. Esto indica que la emisión de la fluorescencia de esta sonda no es tan eficiente, lo cual conlleva a una peor discriminación alélica, tal y como se presencia en la Figura 12.

Tabla 19. Señal esperada y valores medios Cq obtenidos en las curvas de amplificación, y alelos esperados en los gráficos de discriminación alélica; para los seis posibles genotipos del gen CSN3.

| Genotipo | Sondas 13104 FAM_13104_A y HEX_13104_C | | | | Sondas 13124 Cy5_13124_A y FAM_13124_G | | | |
|----------|--|-----------------|-------|-----------------------------|--|-----------------|--------|-----------------------------|
| | Señal esperada ¹ | Cq ² | | Alelo esperado ³ | Señal esperada ¹ | Cq ² | | Alelo esperado ³ |
| | | Media | SD* | | | Media | SD* | |
| AA | FAM | 28,62 | 0,064 | ● | Cy5 | 25,32 | 0,014 | ■ |
| | HEX | - | - | | FAM | - | - | |
| AB | FAM | 29,44 | 0,200 | ▲ | Cy5 | 25,96 | 0,389 | ■ |
| | HEX | 27,42 | 0,012 | | FAM | - | - | |
| BB | FAM | - | - | ■ | Cy5 | 28,93 | 0,281 | ■ |
| | HEX | 27,51 | 0,003 | | FAM | 37,84 | 0,818 | |
| AE | FAM | 27,37 | 0,055 | ● | Cy5 | 25,52 | 0,236 | ▲ |
| | HEX | - | - | | FAM | 35,57 | 0,099 | |
| BE | FAM | 29,36 | 0,192 | ▲ | Cy5 | 29,18 | 0,032 | ▲ |
| | HEX | 27,24 | 0,070 | | FAM | 27,82 | 0,642 | |
| EE | FAM | 26,63 | 0,335 | ● | Cy5 | - | - | |
| | HEX | - | - | | FAM | 13,41 | 16,477 | ● |

¹La señal esperada hace referencia a la señal fluorescente de la sonda que se prevé observar en la curva de amplificación, por lo que también se espera obtener un valor Cq. Las sondas emiten fluorescencia si en la muestra se encuentra su variante específica, puesto que sus secuencias son complementarias. De esta forma, para cada genotipo se indica en sombreado y en negrita, la sonda que es específica y emite señal: FAM_13104_A en los genotipos que contengan los alelos A y/o E, HEX_13104_C en los genotipos compuestos por el alelo B, Cy5_13124_A en los genotipos conformados por los alelos A y/o B, y FAM_13124_G en los genotipos que presenten el alelo E.

²Los valores Cq corresponden a los valores medios obtenidos de las dos réplicas amplificadas por genotipo. El símbolo “-” indica que la señal fue negativa, por lo que en la muestra no se encontraba un genotipo que contuviera alelos específicos de la correspondiente sonda. En rojo se indica si se detectó una señal no esperada, es decir, si emitió fluorescencia la sonda no específica para ese genotipo.

³El alelo esperado hace referencia al alelo que se prevé identificar en el gráfico de discriminación alélica (Figura 12). En la posición 13104, los “●” representan los alelos A y E y los “■” el alelo B. En la posición 13124, los “■” representan los alelos A y B y los “●” el alelo E. De esta forma, los “●” y “■” representan genotipos homocigotos y los “▲” genotipos heterocigotos.

* SD = Standard Deviation.

Por un lado, en una de las réplicas del genotipo EE, la sonda FAM_13124_G no emitió fluorescencia. Además, en la identificación del genotipo heterocigoto AE, la señal de la sonda FAM_13124_G correspondiente al alelo E fue considerada negativa, ya que la intensidad de la señal correspondiente al alelo A emitida por la sonda Cy5_13124_A, fue mucho mayor. Por el contrario, los resultados de las sondas de la posición 13104 fueron óptimos, todos los genotipos

fueron identificados de forma correcta y se observó baja variabilidad entre las réplicas (Figura 12).

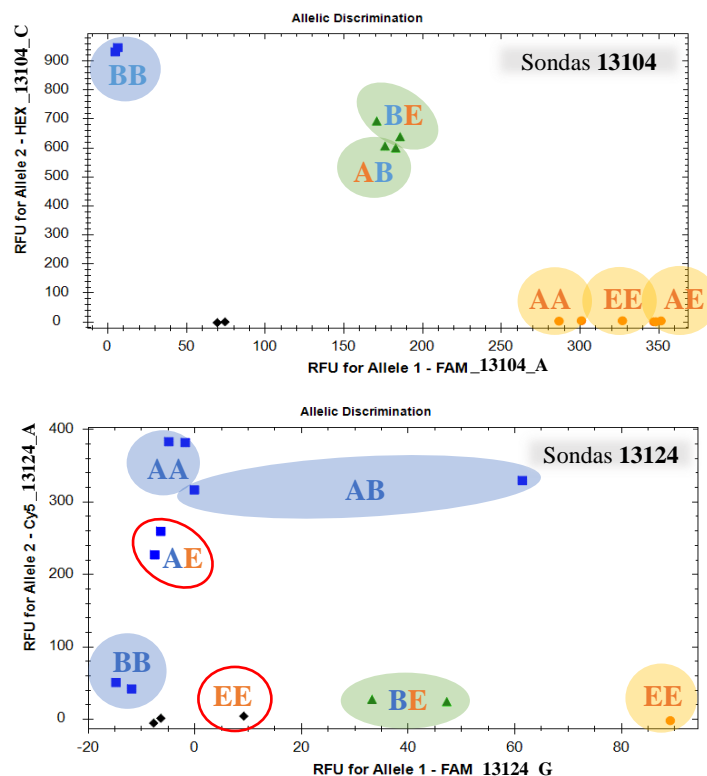


Figura 12. Gráficos de discriminación alélica de las parejas de sondas de la posición 13104 y 13124 (arriba RFU de la sonda HEX_13104_C frente al RFU de la sonda FAM_13104_A, debajo RFU de la sonda Cy5_13124_A frente al RFU de la sonda FAM_13124_G). En los gráficos se asocian las dos réplicas amplificadas de cada genotipo con los alelos identificados por el equipo. En rojo se señalan los genotipos que fueron mal identificados o no identificados, de acuerdo a lo indicado en la Tabla 19. Los “♦” son los puntos no asociados con ningún alelo y corresponden a los blancos.

Finalmente, se concluye que la técnica desarrollada para el genotipado con las sondas de la posición 13104, permite discriminar de forma óptima las variantes A y E, de la variante B. No obstante, en el caso de las sondas de la posición 13124, el funcionamiento no del todo óptimo de la sonda marcada con el fluoróforo FAM y la gran diferencia en la intensidad de su señal fluorescente con la señal de la sonda Cy5_13124_A, supone identificaciones imprecisas en algunos casos. Esto sugiere la necesidad del rediseño de la sonda FAM_13124_G, por ejemplo, mediante el empleo de un fluoróforo diferente, para lograr el funcionamiento equilibrado entre las sondas que permita discriminar el alelo E de los otros dos alelos.

5. CONCLUSIONES

Las conclusiones que pueden extraerse de la técnica desarrollada para el genotipado de ganado bovino de la raza Frisona, son las siguientes:

- La extracción de ADN a partir de las células somáticas de la leche previamente congelada a -80°C durante dos días, ofrece una mayor cantidad y calidad de ADN ($20,99 \text{ ng}/\mu\text{l}$ y $1,92$ de $A_{260/280}$, respectivamente) entre las alternativas ensayadas. Estos resultados sugieren que antes de realizar la extracción, es preferible conservar las muestras a bajas temperaturas, ya que estas podrían favorecer a una mayor liberación de ADN, y por tanto, a una mayor disponibilidad para su extracción.
- Entre los dos tipos de matriz, la extracción a partir de pelo reflejó mayor cantidad y mejor calidad de ADN, con respecto a la mayoría de los grupos muestrales de leche. No obstante, la cantidad extraída a partir de la leche previamente conservada a -80°C , fue significativamente mayor ($10,74 \pm 9,453 \text{ ng}/\mu\text{l}$ o un valor Cq de $26,29 \pm 1,035$ frente a $20,99 \pm 2,562 \text{ ng}/\mu\text{l}$ o un valor Cq de $18,15 \pm 5,231$, para el pelo y leche, respectivamente). Pese a ello, el pelo se considera una matriz de gran interés, ya que se ha observado que la extracción a partir de leche ofrece resultados muy variables debido a la presencia de inhibidores y al contenido variable de células somáticas. De esta forma, perfeccionar la técnica de extracción de ADN a partir de pelo, optimizaría los resultados y permitiría realizar el genotipado en animales de cualquier sexo y edad, lo cual sumaría grandes ventajas a la aplicabilidad de la técnica.
- Las concentraciones de sondas y cebadores con las que se consideró optimizada la técnica fueron $0,2 \mu\text{M}$ y $0,4 \mu\text{M}$, respectivamente. Dichas concentraciones permiten la correcta amplificación de la región de interés del gen *CSN3*, así como la correcta discriminación alélica de las variantes A, B y E, mediante las dos parejas de sondas específicas en reacciones duplex. No obstante, el funcionamiento de la pareja de sondas de la posición 13104, fue notoriamente más equilibrado y eficiente.
- El límite de detección del método se fijó en $6,73 \times 10^0$ copias/ μl para el alelo A, B y E, mientras que los rangos dinámicos definidos, comprendieron de $6,73 \times 10^0$ copias/ μl hasta $6,73 \times 10^2$ o $6,73 \times 10^1$ copias/ μl , dependiendo de la variante a identificar y/o la pareja de sondas empleada. La validez de la técnica se confirmó, al corroborar que en todos los rangos dinámicos la eficiencia presenta valores cercanos al 100% y que la repetibilidad se ajusta a los valores requeridos.
- La técnica desarrollada permite discriminar de forma óptima las variantes A y E, de la variante B, mediante las sondas de la posición 13104. Sin embargo, el funcionamiento no tan óptimo de la sonda FAM_13124_G, impide la correcta discriminación del alelo E de los otros dos alelos, mediante las sondas de la posición 13124. Puesto que para genotipar un individuo se debe amplificar las muestras con las dos parejas de sondas, para garantizar la aplicabilidad de la técnica en duplex es preciso el rediseño de la sonda FAM_13124_G. Entre las posibles soluciones, se valora el uso de un fluoróforo reportero que presente colores más brillantes y permita a la sonda alcanzar mayores valores de RFU, o ensayar diferentes combinaciones fluoróforo-sonda, hasta equilibrar los niveles de fluorescencia de la pareja.

Finalmente, al igual que en trabajos anteriormente publicados, se ha evidenciado la aptitud de la técnica PCR en tiempo real en el genotipado de ganado bovino, ofreciendo una alternativa eficiente a las técnicas tradicionalmente empleadas. Concretamente, la técnica desarrollada

pretende facilitar a los ganaderos la selección genética del ganado bovino de la raza Frisona, orientada a lograr el aumento de la frecuencia alélica de la variante B de la kappa-caseína y la disminución de las variantes A y E. Este sería aplicable especialmente en sistemas que quieran diversificar y transformar la leche a derivados lácteos como queso o yogures. De esta forma, se busca mejorar el rendimiento productivo en las industrias lácteas, lo cual conlleva a la generación de mayores beneficios económicos, a una mayor competencia en el mercado y a la satisfacción de las exigencias del “actual consumidor”. No obstante, como se ha indicado, para ello previamente se debe optimizar el funcionamiento de las sondas específicas en duplex. Por otra parte, por medio de futuros ensayos, también existe la posibilidad de perfeccionar posteriormente la técnica para su desarrollo en multiplex.

BIBLIOGRAFÍA

- Abdel-Azim, G., & Freeman, A. E. (2002). Superiority of QTL-Assisted Selection in Dairy Cattle Breeding Schemes. *Journal of Dairy Science*, 85(7), 1869–1880. [https://doi.org/10.3168/jds.S0022-0302\(02\)74261-6](https://doi.org/10.3168/jds.S0022-0302(02)74261-6)
- Adamov, N., Atanasov, B., Ilievska, K., Nikolovski, M., Dovenska, M., Petkov, V., & Dovenski, T. (2020). Allele and Genotype Frequencies of the Kappa-Casein (CSN3) Locus in Macedonian Holstein-Friesian Cattle. *Macedonian Veterinary Review*, 43(1), 45–54. <https://doi.org/10.2478/macvetrev-2020-0013>
- Al-Khudhair, A., VanRaden, P. M., Null, D. J., & Li, B. (2021). Marker selection and genomic prediction of economically important traits using imputed high-density genotypes for 5 breeds of dairy cattle. *Journal of Dairy Science*, 104(4), 4478–4485. <https://doi.org/https://doi.org/10.3168/jds.2020-19260>
- Alonso López, L., & Mayo Pérez, B. (2015). Productos Lácteos Fermentados y Salud. *Alimentación, Nutrición y Salud*, 22(2), 40–48.
- Bär, C., Mathis, D., Neuhaus, P., Dürr, D., Bisig, W., Egger, L., & Portmann, R. (2019). Protein profile of dairy products: Simultaneous quantification of twenty bovine milk proteins. *International Dairy Journal*, 97, 167–175. <https://doi.org/10.1016/j.idairyj.2019.01.001>
- Barroso, A., Dunner, S., & Cañón, J. (2001). Aplicaciones de las técnicas de la genética molecular para el estudio del polimorfismo genético de las proteínas lácteas de los rumiantes domésticos. Revisión. *ITEA-Información Técnica Económica Agraria*, 97A-I, 11–28.
- Bartonova, P., Vrtkova, I., Kaplanova, K., & Urban, T. (2012). Association between CSN3 and BCO2 gene polymorphisms and milk performance traits in the Czech Fleckvieh cattle breed. *Genetics and Molecular Research*, 11(2), 1058–1063. <https://doi.org/10.4238/2012.April.27.4>
- BIO-RAD Laboratories. (2006). *Real-Time PCR Applications Guide*. <https://bit.ly/3iTcut0>
- BIO-RAD Laboratories. (2021). *Levels of relative fluorescence units (RFUs)*. Retrieved July 3, 2021, from <https://bit.ly/3k53bWd>
- Boichard, D., Boussaha, M., Capitan, A., Rocha, D., Hozé, C., Sanchez, M., ... Guldbrandtsen, B. (2018). Experience from large scale use of the EuroGenomics custom SNP chip in cattle. *Proceedings of the World Congress on Genetics Applied to Livestock Production, Molecular Genetics 4*, 675.
- Brugè, F., Littarru, G. P., Silvestrini, L., Mancuso, T., & Tiano, L. (2009). A novel Real Time PCR strategy to detect SOD3 SNP using LNA probes. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 669(1–2), 80–84. <https://doi.org/10.1016/j.mrfmmm.2009.05.007>
- Carlos Solarte, P., Carol Rosero, G., & Yohana Eraso, C. (2012). Comparación de metodologías moleculares para identificar el gen de la kappa caseína en ganado Holstein. *Revista MVZ Cordoba*, 17(1), 2878–2883. <https://doi.org/10.21897/rmvz.256>
- Caroli, A. M., Chessa, S., & Erhardt, G. J. (2009). Invited review: Milk protein polymorphisms in cattle: Effect on animal breeding and human nutrition. *Journal of Dairy Science*, 92(11), 5335–5352. <https://doi.org/10.3168/jds.2009-2461>
- Chessa, S., Gattolin, S., Cremonesi, P., Soglia, D., Finocchiaro, R., Van Kaam, J. T., ... Civati, G. (2020). The effect of selection on casein genetic polymorphisms and haplotypes in Italian Holstein cattle. *Italian Journal of Animal Science*, 19(1), 833–839. <https://doi.org/10.1080/1828051X.2020.1802356>
- Codex Alimentarius. (2010). *Guidelines on Performance Criteria and Validation of Methods for Detection, Identification and Quantification of Specific DNA Sequences and Specific Proteins in Foods* (CXG 74-2010). <https://bit.ly/3kZrbdF>
- Cole, J. B., Eaglen, S. A. E., Maltecca, C., Mulder, H. A., & Pryce, J. E. (2020). The future of phenomics in dairy cattle breeding. *Animal Frontiers*, 10(2), 37–44. <https://doi.org/10.1093/af/vfaa007>
- Cortés-López, N. G., Del Moral, S., Rueda, J. A., Luna-Palomera, C., Meza-Herrera, C. A., & Abad-Zavaleta, J. (2012). Allelic and genotypic frequency of kappa casein gene in double purpose cattle. *Tropical and Subtropical Agroecosystems*, 15(1), 47–55.
- Duifhuis-Rivera, T., Lemus-Flores, C., Ayala-Valdovinos, M., Sánchez-Chiprés, D. R., Galindo-García, J., Mejía-Martínez, K., & González-Covarrubias, E. (2014). Polymorphisms in beta and kappa-casein are not associated with milk production in two highly technified populations of holstein cattle in Mexico. *Journal of Animal and Plant Sciences*, 24(5), 1316–1321.
- Fonseca, P. A. S., Rosse, I. C., DeMiranda, M., Machado, M. A., Verneque, R. S., Peixoto, M. G. C. D., & Carvalho, M. R. S. (2013). A new tetra-primer ARMS-PCR for genotyping bovine kappa-casein polymorphisms. *Genetics and Molecular Research*, 12(4), 6521–6526.

<https://doi.org/10.4238/2013.December.11.3>

- Fontanesi, L., Calò, D. G., Galimberti, G., Negrini, R., Marino, R., Nardone, A., ... Russo, V. (2014). A candidate gene association study for nine economically important traits in Italian Holstein cattle. *Animal Genetics*, 45(4), 576–580. <https://doi.org/10.1111/age.12164>
- Food and Agriculture Organization of the United Nations (FAO). (2007). *The State of the World's Animal Genetic Resources for Food and Agriculture*. (B. Rischkowsky & D. Pilling, Eds.). Rome.
- Food and Drug Administration (FDA). (2019). Guidelines for the Validation of Analytical Methods for Nucleic Acid Sequence-Based Analysis of Food, Feed, Cosmetics and Veterinary Products. *Foods Program Methods Validation Processes and Guidelines*. <https://bit.ly/3th0sgA>
- Giglioti, R., Gutmanis, G., Katiki, L. M., Okino, C. H., de Sena Oliveira, M. C., & Vercesi Filho, A. E. (2020). New high-sensitive rhAmp method for A1 allele detection in A2 milk samples. *Food Chemistry*, 313(November 2019), 126167. <https://doi.org/10.1016/j.foodchem.2020.126167>
- Goddard, M. E., & Hayes, B. J. (2007). Genomic selection. *Journal of Animal Breeding and Genetics = Zeitschrift Fur Tierzucht Und Zuchtungsbiologie*, 124(6), 323–330. <https://doi.org/10.1111/j.1439-0388.2007.00702.x>
- González-Recio, Ó., Carabaño Luengo, M. J., Pena Alberdi, J., Díaz Martín, C., Ugarte Sagastizabal, E., & Alenda Jiménez, R. (2009). Plan de implantación de la selección genómica en la población de vacuno de leche en España. In *XXXIX Jornadas de Estudio. XIII Jornadas sobre producción animal* (Vol. 1, pp. 123–125). Zaragoza.
- Hayes, B. J., Bowman, P. J., Chamberlain, A. J., & Goddard, M. E. (2009). Invited review: Genomic selection in dairy cattle: Progress and challenges. *Journal of Dairy Science*, 92(2), 433–443. <https://doi.org/https://doi.org/10.3168/jds.2008-1646>
- Integrated DNA Technologies (IDT). (2012). *PrimeTime qPCR Application Guide 3rd Edition*. <https://bit.ly/3iZXvxA>
- Integrated DNA Technologies (IDT). (2021a). *Tips for working with gBlocks Gene Fragments*. Calculating copy number from concentration measurements. Retrieved June 6, 2021, from <https://bit.ly/3D1CjPo>
- Integrated DNA Technologies (IDT). (2021b). *OligoAnalyzer™ Tool*. Retrieved May 3, 2021, from <https://bit.ly/37U8Bh7>
- Integrated DNA Technologies (IDT). (2021c). *PrimerQuest™ Tool*. Retrieved May 3, 2021, from <https://bit.ly/3stwDsF>
- Ionel, N. R., Elena, I. D., & Toma, C. L. (2017). Effects of sires genotyping for k-Casein on B allele frequency in Romanian Brown Cattle. *Research Journal of Biotechnology*, 12(12), 9–13.
- Koopae, H. K., & Koshkoiyeh, A. E. (2014). SNPs genotyping technologies and their applications in farm animals breeding programs: review. *Brazilian Archives of Biology and Technology*, 57(1), 87–95. <https://doi.org/10.1590/S1516-89132014000100013>
- Korhonen, H., & Pihlanto, A. (2006). Bioactive peptides: Production and functionality. *International Dairy Journal*, 16(9), 945–960. <https://doi.org/10.1016/j.idairyj.2005.10.012>
- Kovalchuk, S., Tagmazyan, A., & Klimov, E. (2019). A Novel Test System for Genotyping rs43703016 Single-nucleotide Substitutions in the Bovine CSN3 Gene. *Annual Research & Review in Biology*, 32(4), 1–5. <https://doi.org/10.9734/arrb/2019/v32i430090>
- Latorra, D., Campbell, K., Wolter, A., & Hurley, J. M. (2003). Enhanced allele-specific PCR discrimination in SNP genotyping using 3' locked nucleic acid (LNA) primers. *Human Mutation*, 22(1), 79–85. <https://doi.org/10.1002/humu.10228>
- Macherey-Nagel. (2014). *Genomic DNA from Tissue. User manual, NucleoSpin®Tissue*. <https://bit.ly/3gdvSzj>
- Mahmoudi, P., Rostamzadeh, J., Rashidi, A., Zergani, E., & Razmkabir, M. (2020). A meta-analysis on association between CSN3 gene variants and milk yield and composition in cattle. *Animal Genetics*, 51(3), 369–381. <https://doi.org/10.1111/age.12922>
- Medrano, J. F. (2000). Nuevas tecnologías para el mejoramiento genético de la producción de leche. In *XXI Congreso Mundial de Buatría - XXVIII Jornadas Uruguayas de Buatría* (pp. 1021–1028). Punta del Este, Uruguay: Centro Médico Veterinario de Paysandú.
- Meuwissen, T. H., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157(4), 1819–1829.
- Miluchová, M., Gábor, M., Candrák, J., Trakovická, A., & Candráková, K. (2018). Association of HindIII-polymorphism in kappa-casein gene with milk, fat and protein yield in holstein cattle. *Acta*

- Biochimica Polonica*, 65(3), 403–407. https://doi.org/10.18388/abp.2017_2313
- Ministerio de Agricultura Pesca y Alimentación (MAPAMA). (2021). *Informe del Consumo Alimentario en España 2020. Panel de consumo alimentario*. <https://bit.ly/3sqDiUs>
- Miranda, G., Bianchi, L., Krupova, Z., Trossat, P., & Martin, P. (2020). An improved LC–MS method to profile molecular diversity and quantify the six main bovine milk proteins, including genetic and splicing variants as well as post-translationally modified isoforms. *Food Chemistry: X*, 5(December 2019), 100080. <https://doi.org/10.1016/j.fochx.2020.100080>
- Molee, A., Boonek, L., & Rungsakinnin, N. (2011). The effect of beta and kappa casein genes on milk yield and milk composition in different percentages of Holstein in crossbred dairy cattle. *Animal Science Journal*, 82(4), 512–516. <https://doi.org/10.1111/j.1740-0929.2011.00879.x>
- National Center for Biotechnology Information (NCBI). (2021a). *Basic Local Alignment Search Tool*. Retrieved May 3, 2021, from <https://bit.ly/3j45Rnz>
- National Center for Biotechnology Information (NCBI). (2021b). *CSN3 casein kappa [Bos taurus (cattle)]*. Retrieved May 3, 2021, from <https://bit.ly/3z0gPjK>
- Organisation for Economic Co-operation Development (OECD), & Food and Agricultural Organization of the United Nations (FAO). (2021). *OECD-FAO Agricultural Outlook 2021-2030*. <https://doi.org/https://doi.org/10.1787/19428846-en>
- Peixoto, C. C., Marcelino, I., Vachiéry, N., Bensaid, A., Martinez, D., Carrondo, M. J. T., & Alves, P. M. (2005). Quantification of *Ehrlichia ruminantium* by real time PCR. *Veterinary Microbiology*, 107(3–4), 273–278. <https://doi.org/10.1016/j.vetmic.2005.02.001>
- Pirondini, A., Bonas, U., Maestri, E., Visioli, G., Marmioli, M., & Marmioli, N. (2010). Yield and amplifiability of different DNA extraction procedures for traceability in the dairy food chain. *Food Control*, 21(5), 663–668. <https://doi.org/10.1016/j.foodcont.2009.10.004>
- Pokorska, J., Kułaj, D., Dusza, M., Żychlińska-Buczek, J., & Makulska, J. (2016). New Rapid Method of DNA Isolation from Milk Somatic Cells. *Animal Biotechnology*, 27(2), 113–117. <https://doi.org/10.1080/10495398.2015.1116446>
- Rangel, A. H. N., Zaros, L. G., Lima, T. C., Borba, L. H. F., Novaes, L. P., Mota, L. F. M., & Silva, M. S. (2017). Polymorphism in the Beta Casein Gene and analysis of milk characteristics in Gir and Guzera dairy cattle. *Genetics and Molecular Research*, 16(2). <https://doi.org/10.4238/gmr16029592>
- Requena, F. D., Agüera, E. I., & Requena, F. (2007). Genética de la caseína de la leche en el bovino Frisón. *REDVET. Revista Electrónica de Veterinaria*, VIII(1), 1–9.
- Rojo Martínez, G. (2019). La importancia de los lácteos en la dieta: más allá del hueso. *Nutrición Hospitalaria*, 36(3), 497–498. <https://doi.org/10.20960/nh.02667>
- Sah, B. N. P., Vasiljevic, T., McKechnie, S., & Donkor, O. N. (2018). Antioxidative and antibacterial peptides derived from bovine milk proteins. *Critical Reviews in Food Science and Nutrition*, 58(5), 726–740. <https://doi.org/10.1080/10408398.2016.1217825>
- Sanchez, M. P., Fritz, S., Patry, C., Delacroix-Buchet, A., & Boichard, D. (2020). Frequencies of milk protein variants and haplotypes estimated from genotypes of more than 1 million bulls and cows of 12 French cattle breeds. *Journal of Dairy Science*, 103(10), 9124–9141. <https://doi.org/10.3168/jds.2020-18492>
- Sebastiani, C., Arcangeli, C., Ciullo, M., Torricelli, M., Cinti, G., Fisichella, S., & Biagetti, M. (2020). Frequencies evaluation of β -Casein gene polymorphisms in dairy cows reared in central Italy. *Animals*, 10(2), 1–7. <https://doi.org/10.3390/ani10020252>
- Sellick, G. S. (2002). *Discovery, Analysis, and Utility of Bovine Single Nucleotide Polymorphisms* [Thesis, The University of Adelaide, Dept. of Animal Science]. <https://bit.ly/3khwoxh>
- Spelman, R. J., Coppieters, W., Karim, L., Arendonk, J. A. M. van, & Bovenhuis, H. (1996). Quantitative trait loci analysis for five milk production traits on chromosome six in the Dutch Holstein-Friesian population. *Genetics (USA)*, 144(4), 1799–1808.
- Subdirección General de Producciones Ganaderas y Cinegéticas, & Dirección General de Producciones y Mercados Agrarios. (2021). *Vacuno lechero. Agosto 2021*. (Ministerio de Agricultura Pesca y Alimentación (MAPAMA), Ed.), *Panel Situación Sector Lácteo en España*. Madrid. <https://bit.ly/2W5tsv4>
- Tamay De Dios L, Ibarra, C., & Velasquillo, C. (2013). Pcr medic graphic de la polimerasa (PCR) y de la PCR en tiempo real. *Tecnología En Salud*, 2(2), 70–79.
- Trakovická, A., Moravčíková, N., & Navrátilová, A. (2012). Kappa-casein gene polymorphism (CSN3)

- and its effect on milk production traits. *Acta Fytotechnica et Zootechnica*, 15(3), 1336–9245.
- Tyulkin, S. V., Vafin, R. R., Zagidullin, L. R., Akhmetov, T. M., Petrov, A. N., & Diel, F. (2018). Technological Properties of Milk of Cows with different Genotypes of Kappa-Casein and Beta-Lactoglobulin. *Foods and Raw Materials*, 6(1), 154–162. <https://doi.org/10.21603/2308-4057-2018-1-154-162>
- Ugozzoli, L. A., Latorra, D., Pucket, R., Arar, K., & Hamby, K. (2004). Real-time genotyping with oligonucleotide probes containing locked nucleic acids. *Analytical Biochemistry*, 324(1), 143–152. <https://doi.org/10.1016/j.ab.2003.09.003>
- Usman, T., Yu, Y., Liu, C., Fan, Z., & Wang, Y. (2014). Comparison of methods for high quantity and quality genomic DNA extraction from raw cow milk. *Genetics and Molecular Research*, 13(2), 3319–3328. <https://doi.org/10.4238/2014.April.29.10>
- Vafin, R. R., & Gilmanov, K. K. (2021). Real-Time Pcr Technology for Cattle Genotyping By a and B Kappa-Casein Gene Alleles. *Series Chemistry and Technology*, (February), 126–132. <https://doi.org/10.32014/2021.2518-1491.16>
- Volkandari, S. D., Indriawati, & Margawati, E. T. (2017). Genetic polymorphism of kappa-casein gene in Friesian Holstein: A basic selection of dairy cattle superiority. *Journal of the Indonesian Tropical Animal Agriculture*, 42(4), 213–219. <https://doi.org/10.14710/jitaa.42.4.213-219>
- Vougiouklaki, D., Antonopoulos, D., Allexeli, S., & Houhoula, D. (2020). Identification of Polymorphisms of Gene CSN2 of B Casein in Greek Cow Breeds (Holstein) by Restriction Fragment Length Polymorphism. *Journal of Agricultural Science*, 12(11), 32. <https://doi.org/10.5539/jas.v12n11p32>
- Wassermann, S. (2020). *A1/A2 β -Casein Charakterisierung mittels Real-Time-PCR* [Dissertation, LMU München, Faculty of Veterinary Medicine]. <https://doi.org/https://doi.org/10.5282/edoc.26665>
- Yap, M., Feehily, C., Walsh, C. J., Fenelon, M., Murphy, E. F., McAuliffe, F. M., ... Cotter, P. D. (2020). Evaluation of methods for the reduction of contaminating host reads when performing shotgun metagenomic sequencing of the milk microbiome. *Scientific Reports*, 10(1), 1–11. <https://doi.org/10.1038/s41598-020-78773-6>
- Zambrano Burbano, Gema L, Eraso Cabrera, Yohana M, Solarte Portilla, Carlos E, & Rosero Galindo, C. Y. (2010). Kappa casein genotypes and curd yield in Holstein cows. *Revista Colombiana de Ciencias Pecuarias*, 23(4), 422–428.

ANEXOS

ANEXO 1. Resultados en la optimización de las condiciones de amplificación mediante qPCR.

Optimización del funcionamiento de las sondas

Tabla 20. Resultados individuales para cada sonda en la optimización de su concentración y condiciones de tiempo y temperatura de la fase annealing.

| Prueba N° | Condiciones | Resultados sondas | | | |
|-----------|--|-------------------|------------------|-------------|-------------|
| | | FAM_13104_A | HEX_13104_C | Cy5_13124_A | FAM_13124_G |
| 1 | Pool de ADN Leche -80°C | MAL | MAL | BIEN | MAL |
| | Singleplex | | | | |
| | 0,1 µM y 0,05 µM | | | | |
| 2 | 60°C durante 30'' | MAL | BIEN | BIEN | MAL |
| | Pool de controles sintéticos | | | | |
| | Singleplex | | | | |
| 3 | 0,1 µM y 0,07 µM | MAL | - | - | MAL |
| | 60°C durante 30'' | | | | |
| | Controles sintéticos | | | | |
| 4 | Singleplex | BIEN | MAL | BIEN | MAL |
| | 0,1 µM y 0,07 µM (solo sondas FAM) | | | | |
| | 60°C durante 30'' | | | | |
| 5 | Controles sintéticos | BIEN (mejorable) | BIEN (mejorable) | BIEN | MAL |
| | Duplex | | | | |
| | 0,07 µM | | | | |
| 6 | 60°C durante 30'' | - | - | MAL | MAL |
| | Controles sintéticos | | | | |
| | Duplex | | | | |
| 7 | 0,07 µM | BIEN | BIEN | BIEN | BIEN |
| | 62°C durante 45'' | | | | |
| | Controles sintéticos | | | | |
| | 0,4 µM de cebadores y 0,2 µM de sondas | BIEN | BIEN | BIEN | BIEN |
| | 58°C durante 30'' para sondas de la posición 13104 | | | | |
| | y 60°C durante 30'' sondas de la posición 13124 | | | | |

*En sombreado se destaca el N° de prueba en el que los resultados se consideran óptimos para cada pareja de sondas, para las condiciones de concentración y tiempo y temperatura de annealing fijados en los respectivos ensayos.

Resultados del análisis estadístico en la selección de la concentración de sondas

Tabla 21. ANOVA simple del factor “concentración”, de las sondas HEX_13104_C y Cy5_13124_A.

| | Variables | F | P-valor* |
|-------------|-----------|---------|--------------|
| HEX_13104_C | Cq | 0,360 | 0,581 |
| | End RFU | 163,153 | 0,000 |
| Cy5_13124_A | Cq | 3,519 | 0,134 |
| | End RFU | 9,704 | 0,036 |

*Se destaca en sombreado si existen diferencias significativas (P≤0,05).

Tabla 22. Comparativa de los valores End RFU para la selección de la concentración estadísticamente mejor.

| | Concentración | Media | SD* |
|-------------|----------------|---------------|---------------|
| HEX_13104_C | 0,1 µM | 173,76 | 4,005 |
| | 0,07 µM | 286,82 | 14,800 |
| Cy5_13124_A | 0,1 µM | 360,39 | 11,712 |
| | 0,07 µM | 334,15 | 8,701 |

* SD = Standard Deviation.

**Se destaca en sombreado las concentraciones para cada tipo de sonda que ofrecen mejores valores de RFU asumiendo un error del 5%.

ANEXO 2. Resultados en la validación del método de identificación de alelos A, B y E mediante qPCR.

Cálculo del número de copias de las muestras de referencia para la elaboración de las curvas estándar

$$(C) (M) (1 \times 10^{-15} \text{ mol/fmol}) (\text{Número de Avogadro}) = \text{número de copias}/\mu\text{L}$$

Ecuación 1. Fórmula para el cálculo del número de copias de los fragmentos génicos “gBlocks” (IDT, 2021a), donde C = Concentración (ng/ μ L), M = Peso molecular (fmol/ng) y Número de Avogadro = $6,022 \times 10^{23}$.

$$(10 \text{ ng}/\mu\text{L}) \times (11,18 \text{ fmol/ng}) \times (1 \times 10^{-15}) \times (6,022 \times 10^{23}) = 6,73 \times 10^{10} \text{ copias}/\mu\text{L}$$

Ecuación 2. Cálculo del número de copias presentes en las muestras de referencias de las variantes A, B y E de la kappa-caseína, en base a las especificaciones de IDT dispuestas en la Tabla 6.

Rango dinámico y Eficiencia

Tabla 23. Pendiente, Eficiencia (%) y valor de R^2 de los tres ensayos realizados para las sondas de la posición 13104 (FAM_13104_A y HEX_13104_C). Valores de LOD, LOQ y rangos dinámicos fijados para cada alelo y esta pareja de sondas.

| | Alelo A | | | Alelo B | | | Alelo E | | |
|----------------|---|---------|---------|---|---------|---------|---|---------|---------|
| | Placa 1 | Placa 2 | Placa 3 | Placa 1 | Placa 2 | Placa 3 | Placa 1 | Placa 2 | Placa 3 |
| Pendiente | -3,193 | -3,297 | -3,303 | -3,377 | -3,353 | -3,438 | -3,361 | -3,388 | -3,427 |
| Eficiencia (%) | 105,672 | 101,038 | 100,781 | 97,744 | 98,740 | 95,363 | 98,391 | 97,323 | 95,783 |
| R^2 | 0,987 | 0,993 | 0,981 | 0,998 | 0,999 | 1,000 | 0,996 | 0,999 | 0,995 |
| LOD | 6,73x10 ⁰ copias/ μ l | | | 6,73x10 ⁰ copias/ μ l | | | 6,73x10 ⁰ copias/ μ l | | |
| LOQ | 6,73x10 ¹ copias/ μ l | | | 6,73x10 ² copias/ μ l | | | 6,73x10 ¹ copias/ μ l | | |
| Rango dinámico | 6,73x10 ⁶ a 6,73x10 ¹ copias/ μ l | | | 6,73x10 ⁶ a 6,73x10 ² copias/ μ l | | | 6,73x10 ⁶ a 6,73x10 ¹ copias/ μ l | | |

Tabla 24. Pendiente, Eficiencia (%) y valor de R^2 de los tres ensayos realizados para las sondas de la posición 13124 (Cy5_13124_A y FAM_13124_G). Valores de LOD, LOQ y rangos dinámicos fijados para cada alelo y esta pareja de sondas.

| | Alelo A | | | Alelo B | | | Alelo E | | |
|----------------|---|---------|---------|---|---------|---------|---|---------|---------|
| | Placa 1 | Placa 2 | Placa 3 | Placa 1 | Placa 2 | Placa 3 | Placa 1 | Placa 2 | Placa 3 |
| Pendiente | -3,301 | -3,141 | -3,259 | -3,313 | -3,364 | -3,323 | -2,878* | -3,420 | -3,484 |
| Eficiencia (%) | 100,868 | 108,142 | 102,682 | 100,382 | 98,286 | 99,947 | 122,599* | 96,075 | 93,639 |
| R^2 | 0,982 | 0,981 | 0,970* | 0,999 | 1,000 | 1,000 | 0,968* | 0,997 | 0,996 |
| LOD | 6,73x10 ⁰ copias/ μ l | | | 6,73x10 ⁰ copias/ μ l | | | 6,73x10 ⁰ copias/ μ l | | |
| LOQ | 6,73x10 ¹ copias/ μ l | | | 6,73x10 ¹ copias/ μ l | | | 6,73x10 ² copias/ μ l | | |
| Rango dinámico | 6,73x10 ⁶ a 6,73x10 ¹ copias/ μ l | | | 6,73x10 ⁶ a 6,73x10 ¹ copias/ μ l | | | 6,73x10 ⁶ a 6,73x10 ² copias/ μ l | | |

* Los resultados no se acogen a los requisitos óptimos debido a errores experimentales.

Repetibilidad intra-assay e inter-assay

Tabla 25. Repetibilidad intra-assay de las curvas estándar de las sondas de la posición 13104 (FAM_13104_A y HEX_13104_C), para validar la discriminación alélica de las variantes A, B y E de la kappa-caseína. **

| Sondas 13104 | | Alelo A | | | Alelo B | | | Alelo E | | |
|----------------------|-------|----------------|-------|--------|----------------|-------|--------|----------------|-------|--------|
| ADN copias/μl | Placa | Media Cq | SD* | CV (%) | Media Cq | SD* | CV (%) | Media Cq | SD* | CV (%) |
| 6,73x10 ⁶ | 1 | 16,37 | 0,546 | 3,34 | 15,73 | 0,188 | 1,19 | 15,19 | 0,831 | 5,47 |
| | 2 | 16,86 | 0,215 | 1,28 | 16,12 | 0,054 | 0,34 | 16,54 | 0,294 | 1,78 |
| | 3 | 20,60 | 0,947 | 4,60 | 16,14 | 0,066 | 0,41 | 19,23 | 0,593 | 3,09 |
| 6,73x10 ⁵ | 1 | 20,15 | 0,555 | 2,75 | 19,94 | 0,166 | 0,83 | 20,36 | 0,167 | 0,82 |
| | 2 | 21,28 | 0,019 | 0,09 | 20,33 | 0,026 | 0,13 | 20,80 | 0,046 | 0,22 |
| | 3 | 24,47 | 0,979 | 4,00 | 20,42 | 0,033 | 0,16 | 22,85 | 0,310 | 1,36 |
| 6,73x10 ⁴ | 1 | 22,73 | 0,116 | 0,51 | 22,53 | 0,040 | 0,18 | 22,48 | 0,227 | 1,01 |
| | 2 | 23,50 | 0,123 | 0,52 | 22,72 | 0,026 | 0,12 | 23,07 | 0,134 | 0,58 |
| | 3 | 25,66 | 0,141 | 0,55 | 22,83 | 0,054 | 0,24 | 25,20 | 0,212 | 0,84 |
| 6,73x10 ³ | 1 | 24,34 | 0,262 | 1,08 | 25,39 | 0,146 | 0,57 | 25,47 | 0,169 | 0,66 |
| | 2 | 25,31 | 0,378 | 1,50 | 25,88 | 0,177 | 0,69 | 26,26 | 0,234 | 0,89 |
| | 3 | 27,84 | 0,810 | 2,91 | 26,13 | 0,126 | 0,48 | 28,97 | 0,205 | 0,71 |
| 6,73x10 ² | 1 | 32,11 | 0,332 | 1,03 | 32,66 | 0,094 | 0,29 | 32,00 | 0,463 | 1,45 |
| | 2 | 33,60 | 0,466 | 1,39 | 32,70 | 0,231 | 0,71 | 33,29 | 0,099 | 0,30 |
| | 3 | 35,71 | 0,825 | 2,31 | 32,82 | 0,132 | 0,40 | 35,71 | 0,731 | 2,05 |
| 6,73x10 ¹ | 1 | 32,11 | 0,052 | 0,16 | 32,66 | 0,405 | 1,24 | 32,00 | 0,225 | 0,70 |
| | 2 | 33,60 | 0,352 | 1,05 | 32,70 | 0,439 | 1,34 | 33,29 | 0,162 | 0,49 |
| | 3 | 35,71 | 0,924 | 2,59 | 32,82 | 0,448 | 1,36 | 35,71 | 0,160 | 0,45 |
| 6,73x10 ⁰ | 1 | 33,03 | 0,512 | 1,55 | 36,01 | 2,568 | 7,13 | 33,39 | 0,146 | 0,44 |
| | 2 | 36,32 | 0,453 | 1,25 | 36,07 | 1,034 | 2,87 | 36,63 | 1,116 | 3,05 |
| | 3 | 37,82 | 1,032 | 2,73 | 36,60 | 1,345 | 3,67 | 39,33 | 0,633 | 1,61 |

Tabla 26. Repetibilidad intra-assay de las curvas estándar de las sondas de la posición 13124 (Cy5_13124_A y FAM_13124_G), para validar la discriminación alélica de las variantes A, B y E de la kappa-caseína. **

| Sondas 13124 | | Alelo A | | | Alelo B | | | Alelo E | | |
|----------------------|-------|----------------|-------|--------|----------------|-------|--------|----------------|-------|--------|
| ADN copias/μl | Placa | Media Cq | SD* | CV (%) | Media Cq | SD* | CV (%) | Media Cq | SD* | CV (%) |
| 6,73x10 ⁶ | 1 | 16,53 | 0,114 | 0,69 | 17,26 | 0,081 | 0,47 | 17,02 | 0,186 | 1,09 |
| | 2 | 16,49 | 0,059 | 0,36 | 17,20 | 0,065 | 0,38 | 17,47 | 0,332 | 1,90 |
| | 3 | 15,92 | 0,117 | 0,74 | 16,62 | 0,037 | 0,22 | 17,42 | 0,333 | 1,91 |
| 6,73x10 ⁵ | 1 | 20,95 | 0,043 | 0,21 | 21,54 | 0,056 | 0,26 | 21,24 | 0,202 | 0,95 |
| | 2 | 21,01 | 0,014 | 0,07 | 21,56 | 0,049 | 0,23 | 21,70 | 0,018 | 0,08 |
| | 3 | 20,43 | 0,017 | 0,08 | 21,14 | 0,018 | 0,09 | 21,40 | 0,108 | 0,51 |
| 6,73x10 ⁴ | 1 | 23,20 | 0,079 | 0,34 | 24,00 | 0,068 | 0,29 | 23,46 | 0,281 | 1,20 |
| | 2 | 23,13 | 0,035 | 0,15 | 23,95 | 0,084 | 0,35 | 23,67 | 0,200 | 0,84 |
| | 3 | 24,19 | 2,621 | 10,83 | 23,38 | 0,038 | 0,16 | 23,63 | 0,150 | 0,63 |
| 6,73x10 ³ | 1 | 24,53 | 0,044 | 0,18 | 27,03 | 0,016 | 0,06 | 27,22 | 0,653 | 2,40 |
| | 2 | 24,48 | 0,022 | 0,09 | 26,96 | 0,184 | 0,68 | 27,24 | 0,047 | 0,17 |
| | 3 | 24,11 | 0,016 | 0,07 | 26,49 | 0,070 | 0,26 | 27,31 | 0,003 | 0,01 |
| 6,73x10 ² | 1 | 33,25 | 0,088 | 0,27 | 33,65 | 0,081 | 0,24 | 34,31 | 5,406 | 15,76 |
| | 2 | 32,11 | 0,070 | 0,22 | 34,01 | 0,082 | 0,24 | 33,88 | 0,596 | 1,76 |
| | 3 | 32,73 | 0,061 | 0,19 | 33,17 | 0,110 | 0,33 | 33,63 | 0,610 | 1,81 |
| 6,73x10 ¹ | 1 | 33,25 | 0,134 | 0,40 | 33,65 | 0,144 | 0,43 | 34,31 | 0,629 | 1,83 |
| | 2 | 32,11 | 2,119 | 6,60 | 34,01 | 0,217 | 0,64 | 33,88 | 0,361 | 1,06 |
| | 3 | 32,73 | 0,208 | 0,63 | 33,17 | 0,245 | 0,74 | 33,63 | 0,107 | 0,32 |
| 6,73x10 ⁰ | 1 | 34,11 | 0,061 | 0,18 | 34,25 | 0,089 | 0,26 | 36,89 | 1,590 | 4,31 |
| | 2 | 34,18 | 0,160 | 0,47 | 34,11 | 0,221 | 0,65 | 37,89 | 0,220 | 0,58 |
| | 3 | 33,94 | 0,218 | 0,64 | 34,36 | 0,483 | 1,41 | 37,57 | 0,111 | 0,29 |

Tabla 27. Repetibilidad inter-assay de las curvas estándar de las sondas de la posición 13104 (FAM_13104_A y HEX_13104_C), para validar la discriminación alélica de las variantes A, B y E de la kappa-caseína. **

| Sondas 13104 | Alelo A | | | Alelo B | | | Alelo E | | |
|----------------------|----------------|-------|--------|----------------|-------|--------|----------------|-------|--------|
| | Media Cq | SD* | CV (%) | Media Cq | SD* | CV (%) | Media Cq | SD* | CV (%) |
| 6,73x10 ⁶ | 17,94 | 2,310 | 12,87 | 15,99 | 0,229 | 1,43 | 16,99 | 2,060 | 12,13 |
| 6,73x10 ⁵ | 20,85 | 2,303 | 11,05 | 19,03 | 0,267 | 1,40 | 20,27 | 1,355 | 6,69 |
| 6,73x10 ⁴ | 23,96 | 1,517 | 6,33 | 22,69 | 0,153 | 0,67 | 23,58 | 1,429 | 6,06 |
| 6,73x10 ³ | 25,83 | 1,804 | 6,98 | 25,80 | 0,375 | 1,45 | 26,90 | 1,837 | 6,83 |
| 6,73x10 ² | 30,95 | 1,757 | 5,68 | 29,49 | 0,196 | 0,66 | 30,98 | 1,863 | 6,01 |
| 6,73x10 ¹ | 33,81 | 1,808 | 5,35 | 32,73 | 0,082 | 0,25 | 33,67 | 1,883 | 5,59 |
| 6,73x10 ⁰ | 35,72 | 2,446 | 6,85 | 36,23 | 0,325 | 0,90 | 36,45 | 2,976 | 8,16 |

Tabla 28. Repetibilidad inter-assay de las curvas estándar de las sondas de la posición 13124 (Cy5_13124_A y FAM_13124_G), para validar la discriminación alélica de las variantes A, B y E de la kappa-caseína. **

| Sondas 13124 | Alelo A | | | Alelo B | | | Alelo E | | |
|----------------------|----------------|-------|--------|----------------|-------|--------|----------------|-------|--------|
| | Media Cq | SD* | CV (%) | Media Cq | SD* | CV (%) | Media Cq | SD* | CV (%) |
| 6,73x10 ⁶ | 16,32 | 0,339 | 2,08 | 17,02 | 0,351 | 2,06 | 17,30 | 0,244 | 1,41 |
| 6,73x10 ⁵ | 19,69 | 0,353 | 1,79 | 20,26 | 0,204 | 1,01 | 20,36 | 0,243 | 1,19 |
| 6,73x10 ⁴ | 23,51 | 0,594 | 2,53 | 23,78 | 0,343 | 1,44 | 23,59 | 0,110 | 0,47 |
| 6,73x10 ³ | 24,37 | 0,230 | 0,94 | 26,83 | 0,293 | 1,09 | 27,26 | 0,046 | 0,17 |
| 6,73x10 ² | 29,83 | 0,429 | 1,44 | 30,49 | 0,273 | 0,90 | 30,16 | 1,987 | 6,59 |
| 6,73x10 ¹ | 32,70 | 0,568 | 1,74 | 33,61 | 0,420 | 1,25 | 33,94 | 0,342 | 1,01 |
| 6,73x10 ⁰ | 34,08 | 0,121 | 0,36 | 34,24 | 0,130 | 0,38 | 37,45 | 0,512 | 1,37 |

* SD = Standard Deviation.

**En sombreado se destacan los valores que comprenden el rango dinámico y que fueron empleados para validar la repetibilidad.

ANEXO 3. Resultados de la aplicabilidad de la qPCR en la identificación de los alelos A, B y E de la kappa-caseína.

Selección de muestras genotipadas para evaluar la aplicabilidad

Tabla 29. ANOVA simple para evaluar el efecto del tipo de matriz en la cantidad (ng/μl) y ($A_{260/280}$) calidad de ADN extraído.

| Variables | F | P-valor* |
|---------------|-------|----------|
| ADN (ng/μl) | 2,819 | 0,107 |
| $A_{260/280}$ | 0,013 | 0,909 |

*El análisis se evaluó con el total de muestras ($n=12$).

Puesto que en ambas variables el P-valor es superior a 0,05, se concluye que no existen diferencias significativas y la selección se realiza de acuerdo a lo descrito en la Tabla 30.

Tabla 30. Cantidad (ng/μl) y calidad de ADN extraído ($A_{260/280}$) de las muestras de leche y pelo procedentes de los mismos individuos genotipados.

| Genotipo | ADN (ng/μl) | | | | | | | | $A_{260/280}$ | | | | | | | |
|----------|-----------------------|------|-------|-------|-------------------|-------------|-------|--------|---------------|-------|------|-------|-------------------|-------------|------|-------|
| | Leche | | | | Pelo ² | | | | Leche | | | | Pelo ² | | | |
| | Muestras ¹ | | Media | SD | Muestras | Media | SD | | Muestras | Media | SD | | Muestras | Media | SD | |
| AA | 8190 | 15,2 | 9,10 | 8,627 | 8190 | 9,9 | 8,60 | 1,838 | 8190 | 1,85 | 1,85 | 0,000 | 8190 | 1,87 | 1,82 | 0,078 |
| | 8251 | 3,0 | | | 8189 | 7,3 | | | 8251 | 1,85 | | | 8189 | 1,76 | | |
| AB | 8237 | 4,0 | 3,25 | 1,061 | 8237 | 4,7 | 6,45 | 2,475 | 8237 | 2,15 | 2,08 | 0,106 | 8237 | 1,71 | 1,88 | 0,233 |
| | 8242 | 2,5 | | | 8242 | 8,2 | | | 8242 | 2,00 | | | 8242 | 2,04 | | |
| BB | 8085 | 4,6 | 8,50 | 5,515 | 8085 | 11,1 | 9,60 | 2,121 | 8085 | 1,90 | 1,53 | 0,014 | 8085 | 1,96 | 1,92 | 0,064 |
| | 8087 | 12,4 | | | 8087 | 8,1 | | | 8087 | 1,88 | | | 8087 | 1,87 | | |
| AE | 9092 | 2,0 | 2,45 | 0,636 | 9092 | 12,6 | 7,35 | 7,425 | 9092 | 1,56 | 1,53 | 0,042 | 9092 | 2,01 | 1,96 | 0,071 |
| | 8203 | 2,9 | | | 8203 | 2,1 | | | 8203 | 1,50 | | | 8203 | 1,91 | | |
| BE | 8170 | 7,0 | 4,95 | 2,899 | 8170 | 4,7 | 7,20 | 3,536 | 8170 | 1,85 | 1,97 | 0,163 | 8170 | 1,67 | 1,84 | 0,233 |
| | 9061 | 2,9 | | | 9061 | 9,7 | | | 9061 | 2,08 | | | 9061 | 2,00 | | |
| EE | 9070 | 9,3 | 5,90 | 4,808 | 9070 | 11,4 | 25,25 | 19,587 | 9070 | 1,95 | 2,20 | 0,346 | 9070 | 2,27 | 2,17 | 0,148 |
| | 9031 | 2,5 | | | 9031 | 39,1 | | | 9031 | 2,44 | | | 9031 | 2,06 | | |

¹Los números corresponden a la codificación numérica de los crotales con los que el ganado estaba identificado en la granja S.A.T. Etxeberri. Debido a la disponibilidad limitada de muestras, una de las muestras de pelo y leche del genotipo AA, no procedieron de los mismos individuos (señalado en rojo).

²Entre los dos tipos de muestra procedentes del mismo individuo, se observa de forma general, que en las extracciones de pelo se obtiene una mayor cantidad y calidad de ADN. De esta forma, para evaluar la aplicabilidad se escogen las muestras de pelo, seleccionando una muestra representativa de cada genotipo, y escogiendo la más óptima de acuerdo a lo descrito en el apartado 3.3 de la metodología. Dichas muestras se señalan en negrita.